Running head: DEBIASING CONDITIONAL PREDICTIONS

The Locus of Theory-based Debiasing Effects on Conditional Predictions

Erin Buchanan

University of Mississippi

William S. Maki

University of Arizona

Abstract

Abstract goes here.

The Locus of Theory-based Debiasing Effects on Conditional Predictions

Intro goes here

And here

Method

*Participants*

Participants were recruited from the University of Mississippi human subjects pool and received class credit for their participation. Participants (122 female, 36 male) were all college aged (18-24 yr). There were 84 participants in the Standard instructions group and 74 participants in the Debiasing instructions group.

*Materials*

Four sets of cue-response pairs, *N*=24 each, were selected from the Nelson et al. (2004) association norms; selection was random but subject to the following constraints. Each set was classified as being low or high in forward associative strength and low or high in backward associative strength. For both forward and backward strengths, *x*, low associative strength was defined as 0.00 < *x* ≤ 0.20 and high associative strength was defined 0.50 < *x* ≤ 0.80. The resulting mean low and high forward strengths were 0.06 and 0.63, and the resulting mean low and high backward strengths were 0.06 and 0.64.

These 96 *test pairs* were arranged in a list in sets of 24. Each set contained an equal number (*N*=6) of each type of pair: low-forward/low-backward, high-forward/low-backward, low-forward/high-backward, and high-forward/high-backward. Pairs within the four sets were then randomly ordered. The order of the 96 pairs was reversed to form a second list to control for order effects.

Eight more *debiasing pairs* were selected from the association norms for use in the debiasing instructions. Seven of the pairs varied considerably in forward strength (range 0.01 – 0.54) but were all very high in backwards strength (range 0.75-0.88). The eighth pair was high in forward strength (0.92) and low in backward strength (0.06). An additional three *practice words* (LOST, OLD, ARTICLE) were drawn from the norms for use in illustrating the free association task.

Two booklets were prepared. Each began with an introduction to the concept of association and a short free association list containing the three practice words. The second part of each booklet described the associative rating procedure and presented a sample rating form containing a list of the eight debiasing pairs. The rating instructions were the same as those used in previous studies (Maki, 2007a, 2007b): “Assume 100 college students from around the nation gave responses to each CUE word. How many of these 100 students do you think would have given the RESPONSE word?” Ratings were marked on a 10-point scale (0-9) with each response defined as 0-9, 10-19, …, 80-89, 90-100 (as in Maki, 2007a). The third part of each booklet consisted of a listing of the 96 test pairs with adjacent rating choices (0-9) , 24 per page; the rating scale and rating instructions appeared as reminders at the top of each page. The booklets differed only with respect to the ordering of the 96 test pairs (one the reverse of the other).

*Procedure*

Participants were tested in groups of 15-20. Each group was randomly assigned to receive either the standard or debiasing instructions. Each participant was randomly assigned one of the two booklets. All groups began by reading a short paragraph describing the free association task and then performed the free association demonstration. After writing their responses to the three practice words, the experimenter asked for a show of hands for those responding with FOUND (to LOST; about 75%) and NEW or YOUNG (to OLD; roughly split). Participants were asked to report their responses to ARTICLE to show the variety of answers written. Then, all the groups were introduced to the associative rating task as an alternative way of measuring associations. All the groups used the 10-point scale to rate the eight debiasing pairs. The experimenter then used a show of hands to estimate the groups’ average rating for each pair and engaged the group in a discussion of the ratings. Finally, all the groups rated the 96 test pairs.

The critical manipulation that distinguished the debiasing-instruction groups from the standard-instruction groups occurred during the discussion of the ratings of the eight debiasing pairs. For the standard instruction groups, the experimenter gave some examples of what the correct ratings would be for individual word pairs (such as a 5 or 6 for the pair SISTER-BROTHER which has a forward strength of 0.54). For the debiasing instruction groups, theory-based debiasing instructions modeled after Koriat and Bjork (2006) were used. After the rating of the debiasing pairs, the experimenter explained how word pairs actually have two associative relationships, both forwards and backwards. Examples were given on how a backwards relationship can unduly influence a person’s rating of a forward relationship.  For example, STEAK-SIRLOIN has a very weak forward relationship, *p*(STEAK🡪SIRLOIN) = 0.01, and should only be rated 0 or 1 on the rating scale.  However, STEAK-SIRLOIN has a very strong backwards relationship, *p*(SIRLOIN🡪STEAK) = 0.81, that would likely lead to an inflated rating.  After STEAK-SIRLOIN was explained, both SISTER-BROTHER (high FSG, high BSG) and CHEDDAR-CHEESE (high FSG, low BSG) were discussed to show different combinations of forward and backward strength. Then participants were allowed to complete the 96 ratings on the experimental word-pairs.

Results

For each participant in each group, ratings were averaged across the 24 test pairs in each combination of forward and backward strength and converted to a 100-point scale. The mean ratings are shown in Figure 2. In each panel, the filled symbols represent the results from the low backward association pairs and the open symbols represent the results from the high backward association pairs. The left panel shows the ratings for the Standard instruction group and the right panel shows the ratings for the Debiasing instruction group. Figure 2 shows three principal findings. First, there was a pronounced interaction between forward and backward strength; the slope of the JAM function relating ratings to forward free association probabilities was increased by higher backward associations. Second, the magnitude of that interaction was unchanged by the debiasing instructions. Third, the debiasing instructions did have an effect, namely to reduce ratings overall by about nine points (on the 100-point scale).

The reliability of these observations was tested with a 2 x 2 x 2 mixed within-between analysis of variance using interactions with subjects as error terms.1 The within-subjects factors were Forward (low vs. high) and Backward (low vs. high) associative strength. The between-group factor was Instructions (standard vs. debiasing). The overall effect of Forward strength was significant, *F*(1, 156) = 955.74, η2partial = 0.86, *p* < .001; each of the four Jam functions in Figure 2 show a reliable increase from low to high forward strength, smallest *t* (73) = 14.70*, p* < .001. The overall effect of Backward strength was significant, *F*(1, 156) = 374.10, η2partial = 0.71, *p* < .001; high backwards strength resulted in higher ratings than did low backward strength. The Forward x Backward interaction was also significant, *F*(1, 156) = 69.63, η2partial = 0.31, *p* < .001; the effect of Forward strength on ratings was amplified by Backwards strength.

None of these effects were modulated by instructions. Neither the interaction of Forward strength and Instructions nor the interaction of Backward strength and Instructions were significant, both *F*s < 1. Of particular note, the three-way interaction of Forward Strength, Backward strength, and Instructions was not significant either, *F*(1, 156) = 1.34, η2partial = 0.01.

The effect of the debiasing instructions was confined to an overall decrease in ratings that was independent of the other variables, *F*(1, 156) = 30.92, η2partial = 0.16, *p* < .001. As shown in Figure 2, the intercepts of all the JAM functions were reduced by about the same amount by the debiasing instructions.

Discussion

Lots of talk here

References

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, **12**, 83-87.

Galton, F. (1879). Psychometric experiments. *Brain*, **2**, 149-162.

Garskof, B. E., & Forrester, W. (1966). The relationships between judged similarity, judged association, and normative association. *Psychonomic Science*, **6**, 504-504.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, **103**, 592-596.

Koriat, A. (2008). Alleviating the inflation of conditional predictions. *Organizational Behavior and Human Decision Processes*, **106**, 61-76.

Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 1133-1145.

Koriat, A., Fiedler, K., & Bjork, R. A. (2006). The inflation of conditional predictions. *Journal of Experimental Psychology: General*, **135**, 429-447.

Levin, D.T., Momen, N., Drivdahl, S.B., & Simons, D.J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, **7**, 397-412.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. New York: John Wiley & Sons.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, **15**, 215-233.

Maki, W. S. (2007a). Judgments of associative memory. *Cognitive Psychology*, **54**, 319-353.

Maki, W. S. (2007b). Separating bias and sensitivity in judgments of associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **33**, 231-237.

Maki, W. S. (2008). A database of associative strengths from the strength sampling model: A theory-based supplement to the Nelson, McEvoy, and Schreiber word association norms. *Behavior Research Methods*, **40**, 232-235.

Maki, W. S., & Buchanan, E. (2008). Latent structure in measures of associative, semantic, and thematic knowledge. *Psychonomic Bulletin & Review*, **15**, 598-603.

McFadden, D. L. (2000/2003). Economic choices. In T. Persson (Ed.), *Nobel Lectures, Economic Sciences 1996-2000.* Singapore: World Scientific Publishing.

Nelson, D. L., Dyrdal, G. M., & Goodmon, L. B. (2005). What is preexisting strength? Predicting free association, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review*, **12**, 711-719.

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition*, **28**, 887-899.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, **36**, 402-407.

Sanbonmatsu, D. M., Posavac, S. S., & Stasney, R. (1997). The subjective beliefs underlying probability overestimation. *Journal of Experimental Social Psychology*, **33**, 276-395.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, **34**, 273-286.

Train, K. E. (2003). Discrete Choice Methods with Simulation. New York: Cambridge University Press.

Yellott, J. I. (1977). The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, **15**, 109-144.

Author Note

Erin Buchanan, Department of Psychology, University of Mississippi. William S. Maki, Department of Psychology, University of Arizona.

***We thank so-and-so for terribly destructive comments.***

Correspondence should be addressed to Erin Buchanan, Department of Psychology, University of Mississippi, PO BOX 1848, University, MS, 38677. Email addresses: Erin Buchanan ([embuchan@olemiss.edu](mailto:embuchan@olemiss.edu)) and William S. Maki (wmaki@email.arizona.edu).

Footnotes

1 Analyses of variance using interactions with items (pairs) as error terms yielded the same pattern of significance for all main effects and interactions.