We thank the reviewers for thoughtful feedback. We are grateful that there is a consensus that we have developed (R1) "a very elegant approach for an extremely popular problem with high novelty and broad relevance" and that (R2) "researchers working on pairwise comparisons and preference learning should find this paper to be interesting and valuable." We also appreciate that the reviewers found the main result that it is possible to simultaneously learn an ideal point representation and a metric to be (R4) "interesting, different, and perhaps surprising." Below we respond to the most common critiques raised by the reviewers, indicating changes that we will make to the paper as appropriate. Furthermore, we note that we also plan to make our code available as soon as the review period concludes.

Lack of theoretical guarantees (R1, R2, R3, R4). We are in full agreement that strong theoretical guarantees are extremely important. While our work does not currently have such guarantees, we believe that our algorithm is significant in that it provides strong empirical evidence of the feasibility of simultaneous metric and ideal point recovery. As a starting point, we have explored borrowing ideas from [1], which gives theoretical error bounds on an alternating algorithm for a similar joint learning problem. However, due to the non-linear nature of our observation model, such an extension seems to be non-trivial at this point.

Noise considerations (R2, R3). In our derivation, we pose the problem in a noiseless environment only for simplicity. A similar derivation could be performed by assuming noisy distance estimates (although deriving the algorithm directly from the paired comparison model does not seem straightforward). Moreover, we emphasize that our algorithm can still work when differences in distances are corrupted by additive i.i.d noise. Such noise might arise from an imperfectly learned embedding or as a way of modelling response errors. The slack variables present in the optimization formulation not only increase the stability in the noiseless setting, but also allow for comparison outcomes to be violated due to such noise. Under common paired comparison noise models, we can also replace the loss function in (7) with the negative log-likelihood, provided that this is concave. For example, the loss function for logistic noise, which is equivalent to assuming comparison outcomes follow a Bradley-Terry model, is $\ell(\boldsymbol{d}) = \sum_{k=1}^{P} \log(1 + e^{-y_k(\boldsymbol{Q}_\Gamma \boldsymbol{d})_k})$ for $y_k = +1/-1$. While Bernoulli noise could be considered a simple noise model, we typically consider noise models where more the probability a comparison flips is inversely proportional to $|\boldsymbol{d}_i - \boldsymbol{d}_j|$, hence why we assume additive noise. We will include a more thorough discussion about noise considerations in our revision.

Experiments: dataset choice, comparison with other techniques (R1, R3, R4). We chose to test our algorithm on a graduate admissions dataset instead of traditional recommender system datasets (MovieLens, UT Zappos50K, etc). This was primarily because with traditional recommender datasets, we would have to construct an embedding first prior to preference/metric estimation. The source of error for unexpected outcomes would then be hard to pinpoint, as it could originate from model mismatch, a poorly learned embedding, or our estimation algorithm. Moreover, in such contexts we could not form interpretable hypotheses about which combinations of features should matter, further hindering our ability to evaluate the quality of the learned metric. The graduate admissions dataset alleviates these difficulties.
For similar reasons, we also did not compare our method against algorithms utilizing different models of preference. In our synthetic experiments, our priority was to quantify *estimation errors* in a setting where there was a known ground truth. While a more thorough comparison of our method with other models of preference on real-world data would certainly be interesting, explaining any gaps in performance (which might be due to either model-mismatch, estimation errors within the model, or a combination) would be difficult and would be beyond the scope of our paper.
Finally, R4 also points out that a similar process of implicitly learning a distance metric occurs in matrix factorization techniques. This is true and something we will emphasize more clearly in our revision, but we want to emphasize that this is different from our work in a key respect. Matrix factorization assumes user preference can be captured by (weighted) inner products, which is a *monotonic* function of the item features that assumes more (or less) of any particular feature is always better. This is distinct from our model, which allows for non-monotonic functions to describe user preference, meaning that there is a "sweet spot" for each item attribute.

Practical implementation (R4). As with any recommender system, practical considerations are important. R4 points out that solving a semidefinite program limits the capability of the algorithm to operate at an industry scale. We recognize this bottleneck and will add discussion on the necessity of practical tools, such as tailored SPD solvers or heuristically driven non-SPD approaches for large scale problems. Also mentioned was the idea of aggregating responses from different users and solving for their respective ideal points under the assumption that all users share the same metric. In our algorithm's current form, we cannot simply aggregate responses as each $\boldsymbol{d}$ vector is personalized for a specific user. However, an extension of our work would be incorporate multiple users and the constraints that arise from collecting (quantized) distance measurements between a set of ideal points and items in a fixed embedding. We consider this extension extremely interesting, but highly non-trivial, and defer it to future work.

Finally, we also plan to update the paper based on more detailed critiques, such as how regularization parameters were chosen (R1), or why the ideal point may slightly deviate outside of "acceptable" ranges (R3).

[1] Sheng Chen and Arindam Banerjee. An improved analysis of alternating minimization for structured multi-response regression. In *Advances in Neural Information Processing Systems*, pages 6616–6627, 2018.