



Robust Modeling Through Design Optimization

Mark A. Pitt¹ · Jay I. Myung¹

© Society for Mathematical Psychology 2019

Abstract

Lee et al. (Computational Brain & Behavior <https://doi.org/10.1007/s42113-019-00029-y>, 2019) mention the use of optimal experimental design to improve model robustness. We elaborate on the usefulness of this tool, describing its benefits and limitations.

Keywords Design optimization · Experimental design · Active learning · Cognitive modeling

Lee et al. (2019) thoughtfully outline a number of good practices in cognitive modeling. They briefly mention the use of optimal experimental design to improve model robustness. We expand on the usefulness of this computational tool, describing its benefit and limitations when used for robust modeling.

A goal of cognitive modeling is to improve inference about the brain mechanisms underlying behavior. The practices that Lee et al. (2019) describe are intended to assist in achieving this goal. Productive modeling also depends on good data, which in turn depends on a good experimental design. Unfortunately, not all experimental designs are equally informative, and as researchers well know, it can be difficult to design an experiment that produces clear-cut results. One reason for this difficulty is that the consequences of design decisions are not known in advance of data collection. Heuristic designs based on one's experience in the field and the success of past studies, or pilot experiments, can assist in improving some design choices, but they rarely provide sufficient insight into experiment outcomes.

The power of computational methods used in robust modeling can be extended to improve experimental design. Adaptive design optimization (ADO; Myung et al. 2013) is one such method for improving inference by maximizing the informativeness (precision) of data collected in an experiment. It is a form of

active learning in machine learning (Settles 2012), which is gaining popularity and accessibility due to the wide availability of fast computing and powerful algorithms. Good measurement precision is necessary for robust modeling, and ADO is one tool that attempts to maximize precision.

The ADO algorithm works by using the cognitive model, combined with participant responses collected over the course of the experiment, to guide stimulus selection trial after trial, to attain the experimental objective in an optimal and efficient manner. The stimulus that is chosen on each trial is intended to be most informative or diagnostic with respect to the specific objective, whether parameter estimation or model discrimination, and works by combining on-the-fly analysis of model behavior and participant behavior in past trials. In parameter estimation, ADO chooses stimuli in the design space that are deemed likely to reduce the greatest uncertainty about parameter values in an information theoretic sense. In model discrimination, ADO chooses those stimuli for which the models make the most disparate predictions. Evidence can accrue quickly, making experiments efficient, a feature especially beneficial to high-cost (e.g., time, personnel, and money) research projects, such as those using expensive imaging methods (e.g., fMRI and MEG) or difficult-to-recruit populations (e.g., children and elderly). As a concrete example, precise estimation of the discounting rate parameter in the hyperbolic model of the delay discounting task (Green and Myerson 2004) can be obtained in 10–20 trials or under 1–2 min of testing (Ahn et al. 2019). In a way, ADO is an algorithmic procedure for boosting statistical power without necessarily increasing sample size.

In addition to being used in experiments of model evaluation and comparison, it is good practice to use ADO prior to data

✉ Mark A. Pitt
Pitt.2@osu.edu

¹ Department of Psychology, Ohio State University,
Columbus, OH 43210, USA

collection. An ADO experiment can be simulated using the proposed experimental design. The outcomes would then provide evidence of what is possible to achieve in the experiment in the most optimistic scenario. Just as importantly, such simulations can alert the researcher to a poor design, because there is no need to run the experiment if you know in advance it is likely to fail. This approach is especially useful for model discrimination studies, where extensive model mimicry can be present. By measuring its extent in advance, researchers are in a more informed position on how to proceed: Revise the design space or run the experiment? Inclusion of such simulation data, along with the precise form of the models being compared, in a preregistered experiment make all details of the study fully transparent and can provide compelling evidence for reviewers about the potential contribution of the study.

All computational methods have limitations and ADO is no exception. Most fundamentally, ADO is not robust to model misspecification. In model discrimination, it assumes one of the contenders is the true model. Conclusions can be misleading or biased depending on how well this assumption holds in practice. Because ADO-generated designs are highly tuned to the assumed models under consideration, the further the underlying cognitive process deviates from any one of the contenders, the less useful the ADO-based designs would be in identifying the true model. This laser focus of ADO on evaluating the models at hand may make the data less well suited for evaluating the suitability of other models.

On a more pragmatic level, ADO's focus on optimizing stimulus selection trial after trial to extract the maximum information in the shortest number of trials makes the method brittle. Although this hypersensitivity makes ADO highly desirable for researching individual differences (clinical populations), this same property can compromise robustness if not careful. For example, participants who do not stay on task (e.g., boredom) can throw the (greedy) algorithm off track because ADO updates its beliefs with each response. Relatedly, ADO can hone in quickly on the most difficult types of trials (these are most informative) and present them repeatedly, which can cause participants to fatigue or get frustrated. Like all machine-learning algorithms, ADO is not plug-

and-play. It can require more or less tuning, including creative modifications of the experiment. When tuned well, it works as advertised, providing excellent test-retest reliability (Ahn et al. 2019; Hou et al. 2016), a necessity for robust modeling. Additional details on the pros and cons of the method can be found in Myung et al. (2013).

For those interested in exploring ADO, an open-source Python package (ADOPy; Yang et al. 2019) is available on Github. It contains code for conducting and simulating ADO-based experiments using three experimental tasks: psychometric function estimation, delay discounting, and risky choice.

In sum, good practices in cognitive modeling are tied to good practices in data collection. ADO is a tool that attempts to inform the latter in service of the former. It is unique in its use of cognitive models to drive experimentation and deserves a place in the modeler's toolbox.

References

- Ahn, W.-Y., Gu, H., Shen, Y., Haines, N., Teater, J. E., Myung, J. I., & Pitt, M. A. (2019). Rapid, precise, and reliable phenotyping of delay discounting using Bayesian adaptive design optimization. *Manuscript under review*.
- Green, L., & Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, *130*(5), 769–792.
- Hou, F., Lesmes, L., Kim, W., Gu, H., Pitt, M. A., Myung, J. I., & Lu, Z.-L. (2016). Evaluating the performance of the quick CSF method in detecting contrast sensitivity function changes. *Journal of Vision*, *16*(6), 18, 1–18,19.
- Lee, M. D., Criss, A. H., Devezer, B., Donkin, C., Etz, A., Leite, F. P., et al. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior*. <https://doi.org/10.1007/s42113-019-00029-y>.
- Myung, J. I., Cavagnaro, D. R., & Pitt, M. A. (2013). A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, *57*, 53–67.
- Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *6*(1), 1–114.
- Yang, J., Pitt, M. A., Ahn, W., & Myung, J. I. (2019). *ADOPy: A Python Package for Adaptive Design Optimization*. <https://doi.org/10.31234/osf.io/mdu23>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.