MARKETING SCIENCE
Vol. 28, No. 2, March–April 2009, p. 226
ISSN 0732-2399 | EISSN 1526-548X | 09 | 2802 | 0226

informs

DOI 10.1287/mksc.1080.0476

Commentary

Discussion of the Article "Website Morphing"

Andrew Gelman
Departments of Statistics and Political Science, Columbia University, New York, New York 10027,
gelman@stat.columbia.edu

The article under discussion illustrates the trade-off between optimization and exploration that is fundamental to statistical experimental design. In this discussion, I suggest that the research under discussion could be made even more effective by checking the fit of the model by comparing observed data to replicated data sets simulated from the fitted model.

Key words: discussion; Web; morphing
History: Received: November 17, 2008; accepted: November 17, 2008. Published online in *Articles in Advance* February 24, 2009.

"Website Morphing" (Hauser et al. 2009) is a fun paper to discuss. In particular I like the summary of technical challenges at the beginning: The website must (1) morph based on relatively few clicks, (2) learn which characteristics are best for which customers, (3) use prior information, and (4) be implemented in real time. So often in applied statistics we find ourselves analyzing a data set in isolation or attacking a large problem with limited data; it is refreshing to see such clear goals, and I think I will try to be focused in this way in my future research projects.

At a more technical level, the trade-off between exploitation and exploration is a standard issue in statistical design of experiments: for exploitation (which we call optimization) we want the best solution based on our current state of knowledge; for exploration (in statistical terms, inference) we want the "leverage" that comes from using design points that can be far from practical. The particular tools developed in this article should be useful for other problems such as computerized adaptive testing in educational applications.

My main suggestion for improving this work is to make it more open-ended by incorporating simulation-based model checking, also called posterior predictive checks (as described, for example, in Chapter 6 of our Bayesian statistics text, Gelman et al. 2003). The idea is to use the fitted model to simulate several replicated data sets and then to compare these to the actual data. The comparisons can be graphical (for example, using scatterplots or time-series plots of individual users' patterns of clicks) and numerical (for example, mean squared errors or log likelihoods of observed responses given predicted probabilities from the model). This open-ended exploration gives a framework for the continuous improvement of details of the proposed method as it is being used while keeping the existing inference and optimization structure.

One advantage of a fully Bayesian model (in computer science terminology, a generative model) is that it can be used immediately to simulate replicated data. I recommend that the authors of this article—and the many readers of this journal who will surely want to apply the methods described to their own problems—take advantage of the generative property to explore the ways in which their model fits, and does not fit, reality. Such graphical and numerical diagnostics could be performed using the data gathered from the application of the fitting and optimization procedure to real users.

Acknowledgments
The author thanks the Columbia University Applied Statistics Center, the National Science Foundation, the National Institutes of Health, and Yahoo! Research for partial support of this work.

References

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Rubin. 2003. *Bayesian Data Analysis*, 2nd ed. Chapman and Hall, London.

Hauser, J. R., G. L. Urban, G. Liberali, M. Braun. 2009. Website morphing. *Marketing Sci.* **28**(2) 202–223.