Implementation and dissemination of Bayesian bias mitigation methods

When researchers selectively publish only those studies that achieve statistical significance, average published effect sizes inevitably inflate because the significance threshold acts as a filter; only the studies with the largest effect sizes have low enough p-values to make it through to publication. Studies with smaller, non-significant effects are rarely published, driving up the average effect size. In other words, the statistical significance filter leads researchers to bias their reported effect sizes.

Bakan (1966) remarked that "the very publication practices themselves are part and parcel of the probabilistic processes on which we base our conclusions concerning the nature of psychological phenomena" (p. 427). Only the lucky, tenacious players "win" frequently enough to qualify for extremely competitive grants and faculty positions; players who "lose" will often file those so-called failed studies away and play the statistical significance roulette wheel again (Rosenthal, 1979). I will extend and apply a Bayesian probabilistic model of publication bias to mitigate the effect of this bias on estimates of effect sizes.

I will address psychology's problem of publication bias on two fronts. First, I will develop Bayesian software for formally modeling, identifying, and mitigating the effects of publication bias. I will make this software freely available via GitHub. Second, I will conduct a survey of published reports using this software in an effort to obtain more realistic estimates of effect sizes and assess the amount of bias present in the field.

Bayesian Bias Mitigation Software

The bias mitigation procedure was initially developed by my prospective doctoral adviser Dr. [redacted]. It defines various censoring functions that represent possible generative models of the publication process. These censoring functions index likelihood functions for relevant test statistics. For example, a *no-bias model*, where all studies are published, indexes the typical central and non-central t distributions (for null and non-null effects, respectively). An *extreme-bias model* indexes t distributions that only assign nonzero density in regions where significant results occur (i.e., $|t|>\sim 2$) and nowhere else. The reported test statistics are evaluated under the different likelihood functions to determine how well the models predicted the result, these weights are used to derive each model's respective posterior probability, and a new weighted effect size is obtained by averaging the models' estimates with respect to their posterior probabilities (see [redacted], [redacted] for details). When the various bias models have high posterior probabilities the average effect size will shrink, thus mitigating the nefarious inflationary effect of publication bias.

This bias mitigation process has so far only been implemented in MATLAB code, and in its current form is too computationally expensive for general use. **I propose to implement the software in the free and flexible statistical software R**—with which I have extensive experience—using Stan for the requisite Markov chain Monte Carlo procedures.

Modern statistical software such as R has made Bayesian methods easy to implement for a host of problems, but most psychology researchers are still only exposed to SPSS or other similar commercial products in their training, which provide few, if any, Bayesian analyses. A promising new program (JASP) is currently under development which empowers users to implement a wide range of Bayesian methods. It is free and open-source and features an intuitive user interface. I will integrate meta-analysis methods using the bias mitigation R code into JASP so that it is accessible to researchers not trained to code using R.

Bias Mitigation Survey

I plan to apply this bias mitigation model to results reported in top psychology journals. Since this mitigation model can be applied to any study that results in a univariate test statistic (t, z, F with 1 degree of model freedom, etc.) it can be applied very broadly across different sub-disciplines and experimental designs. I will survey top-tier journals *Psychological Science*, *Journal of Experimental Psychology: General*, and *Cognition* to obtain bias-mitigated estimates of effect sizes on a wide variety of studies. **I predict that there will be similar levels of bias across sub-disciplines, but I predict bias to systematically vary with regard to experimental design. I expect within-subject designs will result in smaller, less frequent bias than between-subject designs.**

Intellectual Merit

Bayesian methods are advantageous over traditional statistics for many reasons. For example, Bayes factors enable quantification of evidence *for* a null hypothesis, accumulating evidence can be monitored with no need to correct for multiple comparisons, and credible intervals have an intuitive evidential interpretation. Bayesian methods can be used to combat publication bias because there are no strict evidence thresholds, it is easy to see when evidence is simply inconclusive, and there is no incentive to hide "null" results because they have informational value in the Bayesian framework. The addition of the bias mitigation procedure to the Bayesian statistical arsenal will help psychologists identify areas of research where the statistical evidence may not be as strong as it appears.

Broader Impacts

To facilitate the adoption of Bayesian statistics, psychology needs outreach from Bayesians sharing resources that are accessible and easy to understand. **I will contribute to the dissemination of Bayesian methods with an invited paper in *Psychonomic Bulletin & Review* and an invited post on the Psychonomic Society's popular blog**. I will provide psychologists with low-threshold introductory resources on Bayesian methods.

Online tutorials, intuitive and free software, and interactive website apps are perfect for making Bayesian ideas accessible to researchers who otherwise would never encounter them. I have written a number of popular tutorial articles on my blog, in a series titled, [redacted]. These entries average over 4,000 unique readers per post and are shared on social media predominantly by psychologists.

Psychologists clearly *want* to learn about Bayesian methods. They know they are valuable but classes don't cover them. **I will create a comprehensive online resource for psychologists to learn what Bayesian statistics is and how it can be applied.** It will cover topics such as Likelihood, Bayes factors, Markov chain Monte Carlo, Publication Bias Mitigation, and others in easily understandable formats. I will incorporate interactive Shiny apps (powered by R) and JASP demonstrations for each topic.

University of California, Irvine is the perfect place to conduct this research. Dr. [redacted] has extensively published on Bayesian methods, and the Cognitive Science department is a world leader in mathematical psychology with a long-standing Bayesian tradition.

References

Bakan, D. (1966). *Psychological Bulletin*, 66(6), 423-437.

[redacted]

Rosenthal, R. (1979). *Psychological Bulletin*, 86(3), 638-641.