Research Statement

Gustavo Diaz
Department of Political Science, McMaster University
diazg2@mcmaster.ca

My research develops standards to navigate the tradeoffs that emerge when one considers research design alternatives before data collection. I use tools from data science and design-based causal inference to identify practices and procedures that researchers can adopt to improve how they approach data at the pre-analysis stage. My current agenda focuses on the design of randomized controlled trials. In the future, I plan to expand to causal inference in quasi-experiments and observational studies, as well as the use of machine learning for statistical inference.

A recurrent goal in the causal inference literature is to minimize bias, or being close to the hypothetical truth on average. Instead, my agenda focuses on optimizing statistical precision, as in producing consistent estimates after multiple realizations of the same data generation process. The literature implicitly assumes that one cam simply improve precision by increasing sample size. However, this is not feasible in most social science applications due to resource considerations.

Even if resources are not an obstacle, even the least intrusive study wastes time and energy from researchers, administrators, and participants. This implies an ethical mandate to identify a research design that maximizes benefits and minimizes harm at the lowest possible cost. An emphasis on statistical precision allows researchers to anticipate and internalize these costs before implementation.

Focusing on randomized controlled trials, this agenda follows two strands. First, researchers often face the choice between alternative experimental designs for which unbiased estimators are already documented. In this case, opting for the design with more precision leads to unforeseen costs in other dimensions. Two projects exemplify how I develop standard to navigate this precision.

In a manuscript forthcoming in the *Journal of Experimental Political Science*, I propose statistical tests to address problems with double list experiments. Social scientists use list experiments in surveys when respondents may not answer truthfully to sensitive questions. When their assumptions are met, list experiments reduce sensitivity biases from misreporting. However, they tend to produce estimates with high variance, which prevents researchers from improving upon direct questioning. Double list experiments promise to remedy this by implementing two parallel list experiments and then aggregating their results, which roughly halves the variance of the estimate for the prevalence of the sensitive trait.

This implies an estimator that is more precise and still unbiased, but their implementation invites questions over its validity. The tests leverage variation in the order in which respondents see the sensitive item to detect whether respondents are reacting to list experiment questions in unintended ways. This provides researchers with a tool to apply this underexplored variant of the technique more widely.

In a paper under review with Erin Rossiter (Notre Dame), we discuss the circumstances under which adopting research design features aimed at improving precision can instead hurt precision through implicit or explicit sample loss. For example, block randomization often improves precision. However, if this requires contacting participants multiple times to collect blocking covariates first and then post-treatment outcomes, then it creates space for attrition that would not exist otherwise, which may offset the precision gains of blocking. Through simulation and the reanalysis of a survey experiment on misinformation in social media, we illustrate how researchers can entertain this precision-retention tradeoff at the pre-analysis stage.

The second reason to focus on statistical precision is that sometimes a more precise yet biased estimator gives more informative answers than an unbiased estimator, which is especially relevant for policy decision-makers. In this case, the researcher is explicitly sacrificing unbiasedness for the sake of precision. Two other projects reflect this issue.

First, in work in progress with Jake Bowers (Illinois) and Christopher Grady (USAID), we discuss the circumstances under which researchers should prefer biased yet more precise estimators in the analysis of experimental data. Once again, an example of this trade-off comes from block-randomized experiments. Block-randomization entails grouping observations in groups or strata, conducting parallel experiments in each, and then calculating a single treatment effect with a weighted average. Previous work in statistics suggests that block-size weights lead to an unbiased estimator of the average treatment effect in this setting. However, because experiments are expensive to implement, we illustrate situations under which researchers may prefer to sacrifice unbiasedness in favor of more precise estimates, which would lead to the use of precision weights instead, the equivalent of using fixed effects in OLS regression. We extend this idea to the use of M-estimator for data with skewed outcomes. We illustrate these ideas through simulation and by revisiting the design of a messaging field experiment conducted by the Office of Evaluation Sciences in the United States.

Second, in a chapter with Christopher Grady (USAID) and Jim Kuklinski (Illinois) in *The SAGE Handbook of Research Methods in Political Science and International Relations*, we discuss the trend of increasingly more complex research designs in survey experimentation. For example, short vignettes are replaced with factorial experiments including many attributes. This increase in complexity allows researchers to isolate causal effects from potential confounders, but raises concerns over whether the tasks that respondents face in survey experiments tap into the intended construct outside of the survey framework. In other words, we draw attention to the trade-off between simple designs that risk confounding bias and complex designs that yield unbiased estimates but risk external validity bias.

I plan to expand my research program toward the use of data science tools to improve statistical inference. An early example comes from my dissertation work on the effect of investigating the use of federal funds among selected mayors in Brazil on the behavior of other mayors in nearby localities. My theory suggests geographic spillovers as a quantity of interest, yet it gives no guidelines for how far away that effect would travel. This presents a tradeoff between operationalizations that introduce bias by being too narrow or too broad. I overcame this problem using a penalized regression framework to choose the optimal upper bound. This innovation received the best poster award in the 2019 Latin American Political Methodology meeting.