Designing Experiments for Better External Validity

Evaluations of education interventions often focus on ‘whether’ programs work and on comparing the point estimates of the impact of various programs to assess their relative effectiveness. While this is no doubt useful, we posit that experimental evaluations can generate more 'generalizable' lessons by (a) paying more attention to theory to design better interventions, (b) conducting theoretically-motivated replications (and tests of heterogeneous effects), and (c) collecting data on intermediate processes to test mechanisms of impact. These issues will be illustrated using examples from experimental evaluations of teacher performance pay programs, the use of contract teachers, the provision of school grants and materials, and school choice programs.