COMMON QUESTIONS AND CONCERNS ABOUT RANDOMIZED EVALUATIONS

J-PAL
ABDUL LATIF JAMEEL POVERTY ACTION LAB
NORTH AMERICA

IS IT ETHICAL TO ASSIGN PEOPLE TO A CONTROL GROUP, POTENTIALLY DENYING THEM ACCESS TO A VALUABLE INTERVENTION?

If there is rigorous evidence that an intervention is effective and sufficient resources are available to serve everyone, it would be unethical to deny some people access to the program. However, in many cases we do not know whether an intervention is effective (it is possible that it could be doing harm), or if there are enough resources to serve everyone. When these conditions exist, a randomized evaluation is not only ethical, but capable of generating evidence to inform the scale-up of effective interventions, or shift resources away from ineffective interventions.

When a program is first being rolled out, or is oversubscribed, financial and logistical constraints may prevent an organization from serving everyone. In such a case, randomization may be a fairer way of choosing who will have access to the program than other selection methods (e.g. first-come, first-served). Conducting a randomized evaluation may change the selection process, but not the number of participants served.

It is also possible to conduct a randomized evaluation without denying access to the intervention. For example, we could randomly select people to receive encouragement to enroll without denying any interested participants access to the intervention. In other cases, it may be useful to compare two different versions of an intervention, such as an existing version and a version with a new component added.

IS IT POSSIBLE TO CONDUCT RANDOMIZED EVALUATIONS AT LOW-COST WITHOUT HAVING TO WAIT YEARS FOR THE RESULTS?

Collecting original survey data is often the most expensive part of an evaluation, but it is not unique to randomized evaluations. Likewise, it is increasingly possible to conduct evaluations at relatively low-cost by measuring outcomes using existing administrative data, instead of collecting survey data.

The length of time required to measure the impact of an intervention largely depends on the outcomes of interest. For example, long-term outcomes for an educational intervention (e.g. earnings and employment) require a lengthier study than shorter-term outcomes, such as test scores, which can be obtained from administrative records.

Finally, the time and expense of conducting a randomized evaluation should be balanced against the value of the evidence produced and the long-term costs of continuing to implement an intervention without understanding its effectiveness.

CAN A RANDOMIZED EVALUATION TELL US NOT JUST WHETHER AN INTERVENTION WORKED, BUT ALSO HOW AND WHY?

When designed and implemented correctly, randomized evaluations can not only tell us whether an intervention was effective, but also answer a number of other policy-relevant questions. For example, a randomized evaluation can test different versions of an intervention to help determine which components are necessary for it to be effective, provide information on intermediate outcomes in order to test an intervention's theory of change, and compare the effect of an intervention on different subgroups.

However, as with any single study, a randomized evaluation is just one piece in a larger puzzle. By combining the results of one or more randomized evaluations with economic theory, descriptive evidence, and local knowledge, we can gain a richer understanding of an intervention's impact.

ARE THE RESULTS OF RANDOMIZED EVALUATIONS GENERALIZABLE TO OTHER CONTEXTS?

The problem of generalizability is common to any impact evaluation that tests a specific intervention in a specific context. Properly designed and implemented randomized evaluations have the distinct advantage over other impact evaluation methods of ensuring that the estimate of an intervention's impact in its original context is unbiased.

Further, it is possible to design randomized evaluations to address generalizability. Randomized evaluations may test an intervention across different contexts, or test the replication of an evidence-based intervention in a new context. Combining a theory of change that describes the conditions necessary for an intervention to be successful with local knowledge of the conditions in each new context can also inform the replicability of an intervention and the development of more generalized policy lessons.