

# ISSUES IN MOUNTING RANDOMIZED EXPERIMENTS IN EDUCATIONAL RESEARCH AND EVALUATION

---

F.V. FALAYE

## ABSTRACT

As the field of education becomes increasingly complex and demanding, researchers more than ever before need to work at the frontiers of knowledge in order to understand the current global issues and problems of education. This calls for rigorous research whose findings are to inform government decision and policy. Also, the effectiveness of any educational intervention needs to be evaluated to know what works and what does not work. The ultimate goal is to re-shape the system of education appropriately. In examining the use of randomized experiments in educational research and evaluation, this paper highlights its merits, methodological and ethical issues, and the problems involved in using randomized experiments in the school setting. It suggests the need for researchers to update their skills in the use of randomized experiments through capacity building to ensure that findings from educational research and evaluation are creditable and defensible.

**KEY WORDS:** Randomized experiment; Educational research; Evaluation; Research ethics; Methodological issues.

## INTRODUCTION

No doubt education is one of the most diverse sectors in any country of the world. As educational sector grows so also its system becomes increasingly complex. Educational system covers a wide range of components, which include the process of teaching/learning, learning materials, the teacher and the learner, achievement and its measurement, finance, and the ultimate learning outcomes, all of which have constituted a monumental challenge to stakeholders in the field education.

As the system grows, the context in which it operates is also changing. For instance, to address the contemporary global demands, the content of education as well as its practice must be updated to accommodate the changes. And to do this educators require new approaches to redirect learning in effective ways. Knowledge is needed on many issues such as organization of schools and classrooms for effective instruction and learning, effective strategies for teaching and the overall preparation of learners for societal challenges.

All of these call for rigorous research in education, because catching up with today's ever changing social, economic and technological environment, educator must not rely only on tools of yesteryears as they may be inadequate for meeting the challenges of today (Bajah, 1995).

Also, research findings whose ultimate goal is to aid educational practice, inform government decision and policy must be evaluated. Evaluation in this light is primarily aimed at finding out what works, what works better and what does not work. Put in another way, the relative effectiveness of any innovation or reformation either in social, health or education sector is judged by evaluation. The outcomes direct the line of action whether to incorporate the new programme if it works or to discard it. Even if new programmes are adopted, an understanding of how they are run is necessary for improvement and documentation of lessons learnt for future replication.

The need to provide credible evidence about, which programme works in education, and indeed other sectors has underscored the use of randomized experiments or trials. However, it is

---

**F.V. Falaye**, Institute of Education, University of Ibadan, Ibadan, Nigeria.

necessary to point out that the adoption of randomized experiments is fairly new, even in health care research where the use of randomized experiments has become a standard practice. According to Boruch (1997), it was not until 1990s when randomized trials started gaining popularity in the medical, social and behavioral sciences, and education. For instance in education, until 1970s, no controlled- field experiments of any scale appear to have been run to understand the effects of standardized testing on students in any country (Boruch, Foley & Grimshaw, 2002.)

As robust as randomized experiments are in discovering what works, they have been slow to get the attention of education practitioners as well as those in the field of social sciences when compared with the physical sciences. Boruch (2002) observed that while researchers have long adopted the randomized field trials to evaluate the effectiveness of medical procedures, educators have finally started to adopt the method.

Given the need to provide evidence on which programmes and policies are effective in improving student's learning outcomes, the use of randomized trials/experiments no doubt will increase in the field of education. But the questions are: why has randomized experiment not been popularly embraced in education? What are the methodological and ethical issues involved in its use? What are the conditions required for mounting randomized experiments? These are the concerns of this paper.

### **Definition of Concepts**

Randomized controlled trial is a term often used interchangeably with randomized control experiment or test, randomized field trial, randomized social experiment or randomized clinical trial. In this type of study, the individual entities are randomly assigned to different interventions in order to test which one works better. When properly conducted, it yields statistically unbiased estimates of the relative effects of behavioural, economic, medical and social interventions which non-randomized quasi-experiments may not give (Boruch, 2003).

Randomized trials, according to Boruch (2002), are a sturdy method of generating defensible evidence about the relative effectiveness of various interventions. In its simplest usage randomized experiment randomly assigns individuals to one or more treatment groups. The purpose is to determine which treatment works better relative to the others.

There are variations to this basic concept of randomization. In the simplest kind of experiment one treatment is assigned to a randomly selected group known as the treatment group, while the control group is not exposed to the treatment. A variety of treatments rather than only one can also be tested. Similarly, in some randomized trials, entire schools or classrooms, rather than individuals are randomly assigned to treatment and control groups. This is referred to as place – based or cluster randomization. A cluster randomized trial is one in which intact social units, or clusters of individuals are randomized to different intervention groups (Donner & Klar, 2000).

In some experiments, a control group may be left out entirely. Instead, the differences in the effects of different treatments mounted in the experiment are measured (Burtless, 2002). The fundamental assumption of the random allocation is that the groups being compared are equivalent. This way, the observed differences in the outcome can be explained clearly by the treatments to which the groups are exposed. Where groups being compared are different, the difference would affect the response of the groups to the treatment(s) given, and hence masks the outcome of the treatments. Such results become unreliable. This attribute of randomization makes it's stronger than other approaches such as quasi-experiments and correlational studies (Keppel & Zedeck, 2002).

Boruch (1997) distinguishes among randomized trials, observational studies and quasi-experiments drawing out its relative merits. Observational studies do not provide the opportunity to assign individuals to treatment groups in accordance with randomization rules. Similarly, quasi-experiments are different from randomized tests. Quasi-experiments depend on other methods such as analysis of covariance and matching methods to control for initial group differences. In practical sense, it is difficult to get two groups that could be matched on the basis of their similar characteristics. Whereas in 'true' experiments, randomization is employed to control for group differences that may influence treatment outcome (Cochran, 1983; Keppel & Zedeck, 2002).

In a simple term, Burtless (2002) defines a randomized field trial otherwise known as social experiment as a controlled experiment that takes place outside a laboratory setting, in the usual environment where social, economic, or educational interactions occur. Burtless's definition reminds us that randomized field trials

cannot be taken out of their social milieu, a characteristic that distinguishes it from controlled experiments that are mounted in the laboratory, typical of the physical sciences. This distinction constitutes a prominent methodological issue involved in randomized trials in educational research, which is discussed in other section of this paper.

Till date, opinions differ among researchers as to whether or not controlled experiments can be mounted in all settings (Boruch, 2002; Gay & Airasian, 2000). Although observational studies, quasi-experiments and other approaches cannot always provide defensible evidence, but they can generate persuasive evidence about the relative effectiveness of programme interventions, which could at least form baseline information for programme implementation. Perhaps it is for this reason that some evaluators are calling for mixed approach to programme evaluation. The next section discusses the advantages of randomized experiments/trials over non-randomized trials in evaluating programme effectiveness.

### Merits of Randomized Trials

There has been considerable discussion as to the relative advantages of randomized trials over non-randomized ones. Randomization cancels selection bias; hence randomized trials are generally regarded as the 'gold standard' for the evaluation of intervention programmes (Donner & Klar, 2000). Because the individuals are randomly assigned to alternative treatments, the effects of the treatments on behaviour can be measured with high reliability.

A good quality randomized trial produces statistically defensible estimates of the effect of the intervention. It allows the researcher and the policy maker to observe the changes brought about by the intervention. Therefore, the results are more likely to have more credibility than those obtained from non-randomized trials.

A unique advantage of random assignment according to the 2003 U.S. Institute of Education guide is that it is possible to evaluate whether the intervention itself, as opposed to other factors, causes the observed outcome. Thus, it is apt in establishing cause and effect relationships. The process of randomly assigning a large group of people to either an intervention group or to a control group increases the level of confidence that there are no systematic differences between the groups. If the trials are properly executed, the resulting difference in outcomes between the intervention

and control groups can confidently be attributed to the intervention and not to other factors. Randomized experiments show the direction of causality between treatment and outcome, which is not easy to establish in non-experimental studies.

Randomized experiments are relatively simpler than non-randomized experiments in terms of implementation and reporting. The methodology is described in simple and straight forward language, and the findings are presented in concise and understandable manner. This simplicity in presenting findings is an advantage over non-experimental studies where results are described using lengthy and sometimes confusing rhetorics.

### Issues Involved in Randomized Experiments

In this section, two issues – methodological and ethical that are germane to randomized experiments are discussed.

**1. Methodological issues:** These are issues related to how and when to set up good quality randomized trials. The conditions and the *modus operandi* of setting up randomized trials have already been discussed. Boruch's (1997) practical guide on randomized experiments for planning and evaluation provides comprehensive information on how to execute randomized trials.

**2. Ethical issues:** Among the variety of criteria often used to assess the appropriateness of randomized trials, this section focuses on the following:

a) **Fairness in the assignment of individuals to treatment groups:** Fairness in this regard has to do with participants having equal chance of being assigned to any of the study groups. To ensure fairness, eligibility to participate in an experiment is better determined before randomization. This ensures that eligible individuals would not be put at a disadvantage relative to others.

b) **Potential benefits of research to participants:** For instance, are participants assigned to beneficial treatment on the basis of favoritism? In experiments, usually the control group does not undergo new interventions. In order to ensure that those in the control group without intervention are not disadvantaged or denied the benefits that participants in the treatment group are provided, every eligible

participant must be given equal opportunity of being assigned to the beneficial treatment group.

c) **Assurance of privacy and confidentiality:** Maintenance of confidentiality of individuals assigned to treatments, and information is so crucial that special laws have to be enacted to protect the privacy of participants in the experiment. This is a common practice in developed nations. Even though assurance of confidentiality is most often stated, it is not often guaranteed in researches carried out in many developing countries. The issue of confidentiality is viewed more seriously in medical and criminal justice research in which loss of confidentiality of information and privacy may lead to stigmatization of individuals involved in the experiment.

d) **Protection from Harm:** There must be assurance that participants are not harmed in any way, either physically or mentally. The research must protect the participants from procedures that can expose them to danger. If any risk exists, participants must be informed. Research procedures that are likely to cause serious harm to participants are oftentimes not allowed, except the research is of great potential benefit. In such a case, informed consent must be obtained from each participant.

e) **Informed Consent:** Agreement to participate in an experiment needs to be sought from each prospective research participant and the consent must be documented. Given informed consent, participants are more likely to remain in the experiment. Some consent form or paper merely requires prospective participants to signify their consent without providing details about the experiment while some spell out clearly some pertinent information such as the purpose of the research, the benefits, the duration, assurance of confidentiality and so on. In the school setting, parental consent needs to be sought before their children or wards are involved in any experiment, particularly if participants are not of age.

### Problems of Randomized Experiments in Education

Although randomized trials are more or less a norm in medical and pharmaceutical sciences, they are not so common in education. For instance, Mosteller and Boruch (2001), hand searched the America Educational Research

Journal for trials in mathematics and science education between 1964 and 1999. Of a total of 1200 articles published in the journal during the period, less than 40 articles focused on randomized trials. A similar search was conducted focusing on randomized trials in Nigeria, and it was discovered that one area where randomized experiments have been utilized mostly is medicine, followed by agriculture. There was no reported publication on randomized experiments in the field of education.

While randomized experiments tend to be considered the optimum approach to evaluating project impact in other fields, it is not so in education due to certain problems which are discussed below.

**1. Funding:** Worldwide, there are many sectors that compete seriously with education for funding support. While research in medical and pharmaceutical sciences will attract government funding, the same is not true of educational research. This situation is worrisome in poor countries where little or no attention is given to educational research.

Funding of educational research is also dependent on the political ideology of any given nation. In the developing countries, different political parties present different manifestoes. Where education is not seen as a cardinal point, there is not likely to be a strong political will to commit a lot of money to research in education.

**2. Time:** Randomized experiments are rigorous in design and their execution takes more time, more so that many units are involved in most educational reforms. For example, the sample may involve schools within cities, states, and countries. For lack of or inadequate funding, there could be a long time lag between the time a randomized experiment is conceived and the period of its implementation.

**3. Nature of education:** Education, according to Bamisaye (1992), is not crises oriented; hence it does not get as much attention as other fields such as medicine and agriculture. Medical problems are most often life threatening, hence it becomes more or less mandatory to mount randomized experiments to find out what works. Inadequate funds hardly hinder such trials. Most often, experiments in medical and pharmaceutical fields are supported by private and non-governmental organizations. The costs are built into the market price of the particular drugs under trial. Invariably, the consumers bear

the costs of conducting the research. This is not the case with education.

**4. Societal Attitude to Research and Evaluation:** This is a serious challenge particularly in developing countries. Research works are most often not taken seriously, while evaluation is seen as a threat. This can be inferred from, first, the government, in terms of poor funding of educational research; second, weak support from administrators, teachers and the students alike. These groups may create a hostile environment for the researchers. Random assignment of individual students, classrooms or schools requires consent from school's authority. Most often they are unwilling to co-operate for certain reasons. Some see researchers as professionally incompetent to evaluate their work. Another problem is skepticism, as people may not be convinced about the outcomes derived from such experiments. To many, external evaluation is judgmental, while to some the exercise is disruptive.

**5. Nature of Randomized Experiments/Trials:** Boruch (2002) highlights the complexity of randomized field trials in education *vis a vis* other fields such as medicine when he concluded based on the fact that the two disciplines operate in different settings, which influence their research approaches. For instance, isolating the effects of a given intervention from all the other influences may present a difficult task in education where many factors can influence the outcome of a randomized trial. For example, the effectiveness of an instructional strategy on students' learning outcomes may depend on several factors outside of the new strategy being tested. Factors of school, home environment, teacher quality and supply may interact with an instructional strategy to affect learners' achievement. What measurements to use in determining instructional effectiveness therefore pose a problem that research design must tackle to ensure valid outcome, whereas in medicine, the researcher has a greater control over the research setting.

**6. Problem of Random Selection and Assignment:** The basic requirements of an experimental design are that participants be selected in an appropriate manner and randomly assigned to the different experimental conditions. In the school setting, it may be difficult to attain random assignment in the true sense of the word. This is because students are in already existing

classes. Therefore, the researcher rather than assign individual students into treatment groups, intact classes are most often selected and assigned to different experimental groups in what is known as quasi-experimental research. The degree of confidence in making inferences of causality in this design is lower than in true experiment design.

**7. Contamination:** This is a methodological problem of randomized experiment in education. In which case, individuals assigned to the groups at the onset of the experiment may change certain characteristics that qualified them for inclusion in the experiment as it progresses, thus invalidating the results. It is also possible for individuals to move in and out of treatment or control group depending on which is seen as more beneficial, since human beings are bound to interact. And in doing so, information is shared on which treatments are more beneficial. More so, it is believed that social phenomena are difficult to subject to experimentation as it is done in research laboratories. Contamination of experiments is, therefore, highly probable.

**8. Ethical problems:** Randomized experiments may be unethical if the results of such trials are not used. Although the government cannot be forced to use them, the question that arises is how ethical it is to allocate scarce resources in the face of competing needs, and to commit so much time in the design, data collection and analysis for the results to be unused? This could result in serious disillusionment. Another ethical problem has to do with the fact that in randomized experiments, some eligible participants may be exposed to treatment that are beneficial while others are denied the benefits. Health care research has resorted to the use of double blindness, a scenario in which neither the subject nor the person administering the treatment knows what is being done.

#### WAY FORWARD

Despite the numerous problems involved in the use of randomized trials in educational research, if seen as challenges that must be surmounted by researchers, no doubt the use of randomized experiments to evaluate the effectiveness of educational policies and programmes will increase.

Persuasive advocacy is central to gaining government support in terms of financing researches in education, and enhancing political will. The policy makers and school administrators

in particular need to be continuously encouraged to cooperate in providing conducive environment when mounting randomized experiments in schools, and to be persuaded that the outcomes will be beneficial to all.

Capacity building and improvement are also essential. Researchers need to update their skills in the 'art' of mounting randomized trials for determining the effectiveness of new educational policies on a small scale before expanding into larger communities. For programmes that are already being implemented, assessing their impact is crucial in order to highlighting the lessons learned for redefining and generating new framework for future programming. Ethical issues that deal with how research ought to be carried out are essential. It therefore behooves the research community to ensure that ethical standards are set. Experiments must be vetted by appropriate units set up, and such units must be empowered to check compliance.

Disseminating information on studies of effects of educational interventions keeps the policy makers and educators abreast of what works better in improving educational outcomes in the learner. It also provides opportunity for researchers in education to clear the ambiguity in the terminologies most often used.

## CONCLUSION

Randomized controlled experiments have produced dependable evidence about effectiveness of interventions in many fields such as medicine, welfare policy, crime and justice and other areas. However, controlled experiments are seldom used to evaluate educational policy. Attempts have been made to review the methodological and ethical issues involved in the use of randomization to assess educational reforms.

While randomized experiments tend to be widely embraced as the best approach to estimating project impact, there are many setbacks that undermine the feasibility of randomized trials as an evaluation strategy in educational research. Methodological problems, funding, political consideration are the major obstacles.

The problem of semantics also makes tracking randomized trials in education difficult, in that many educational researches conducted in the developing nations and titled quasi-experiments, intact classes are in fact randomly assigned to both the intervention and control groups.

Notwithstanding the problems, the use of randomized experiments will increase in educational research as long as concrete evidence is essential to justify the choice of what works best amongst alternative and competing options. Therefore researchers need to sharpen their skills in mounting randomized experimental designs in order to improve the credibility and acceptability of their research findings.

## ACKNOWLEDGMENTS

The author acknowledges the following people and organizations:

The University of Ibadan management for vigorously pursuing the institution's vision of capacity building, thus providing the linkage with the Mac Arthur Foundation; The John D. and Catherine T. Mac Arthur Foundation for providing funding support, and Dr. Robert, F. Boruch, my host at the Graduate School of Education, University of Pennsylvania, Philadelphia, USA.

## REFERENCE

- Bajah, S. T., 1995. 'Education and Politics. Their Interplay as stabilizing factors in the Nigerian Educational System. A paper presented at the National Conference on Stabilizing the Nigerian Educational System organized by the school of Education of the Federal College of Education, Abeokuta.
- Bamisaye, R., 1992. *Sociological Foundations of Nigerian Education*. AMD Publishers Agodi, Ibadan.
- Boruch, R. F., 1997. *Randomized Experiments for Planning and Evaluation: A Practical Guide*. Thousand Oaks, C.A Sage.
- Boruch, R. F., 2002. 'The Virtue of Randomness. In *Feature www. Educationnext.org* Fall.
- Boruch, R. F., 2003. 'Randomized Field Trials in Education'. In *International Handbook of Evaluation*, Kellaghan, T. and Stufflebeam, D. L. (eds) Kluwer Academic Publishers. Great Britain.
- Boruch, R.F., De Monya, D. and Snyder, B., 2002. 'The Importance of Randomized Field Trials in Education and Related Areas. In *Evidence Matters: Randomized Trials in Education Research*, Mosteller, F. and

- Boruch, R. (edts). Brookings Institution Press, Washington D.C.
- Boruch, R.F., Foley, E., and Grimshaw, J., 2002. Estimating the Effects of Interventions in Multiple Sites and Settings: Place-based Randomized Trials, in *Socialvetensk aplig tidskrift nr 9(2-3)*.
- Burtless, G., 2002. Randomized Field Trials for Policy Evaluation: Why Not in Education. In *Evidence Matters: Randomized Trials in Education Research*, Mosteller, F. and Boruch, R.(edts). Brookings Institution Press, Washington D.C.
- Cochran, W., 1983. Planning and Analysis of Observational Studies, Moses, L. and Mosteller, F.( edts). John Wiley and Sons.
- Donner, A. and Klar, N., 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold, Great Britain.
- Gay, L.R and Airasian, P., 2000. Educational Research Competencies for Analysis and Application. Merrill, N.J, USA.
- National Center for Educational Evaluation and Regional Assistance, 2003. Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide.
- Keppel, G. and Zedeck, S., 2002. Data Analysis for Research Designs. W.H Freeman and Co. USA.