

## REVIEW OF EXPERIMENTAL DESIGNS AND METHODS IN ECONOMICS OF EDUCATION RESEARCH

### EĐİTİM EKONOMİSİ ARAřTIRMALARINDA DENEYSSEL TASARIM VE YÖNTEMLERİN GÖZDEN GEÇİRİLMESİ

Huseyin Atakan KESKİN\*   
Mehmet Levent YILMAZ\*\* 

#### Abstract

The field of economics of education has been receiving constant attention with the advent of growth theories stating how education can produce sustainable long-run economic growth and increase people's skills (Hanushek & Woessmann, 2015; Mincer, 1974). Despite its universal relevance, educational research has not been placed on the same scale with fields such as medicine due to seldom usage of robust quantitative research and the dearth of causal inferences (Creemers, Kyriakidēs, & Sammons, 2010). Given this, leading countries in educational research such as the United States (Scimago Journal & Country Rank, 2018) have been initiating "repeated calls for education policy to rely on a foundation of scientifically based research" (Angrist, 2003, para. 1) to nudge the field of education towards using rigorous and innovative methodological methods and experiments (Murnane & Willett, 2011). Using an experimental method is most suitable when the research aims to test the impact of intervention within the respected field of research (Beach & Pedersen, 2016). Henceforward, this paper addresses the opportunities and challenges of using experimental methods in educational interventions, particularly randomized control trials (RCT) and quasi-experiments that test the impact of financial incentives to increase student outcomes. The first section is an overview of experimental designs, followed by sections delineating on RCTs and quasi-experiments, and discussing empirical studies that employ such methods. It should be noted that this paper argues in favor of neither quantitative nor qualitative research methods as both methods can produce quality research if implemented rigorously (Lodico, Spaulding, & Voegtle, 2010).

**Keywords:** Economics of Education; Research Methods; Experimental Research Design; Educational Outcomes; Randomized Controlled Trials; Quasi Experiments; Financial Incentives

**JEL Codes:** I21, C80, C90

\* MSc, University of Oxford, St. Antony's College, Department of Education atakan@fairoopportunityproject.org

\*\* PhD, Turkish National Police Academy, mehmetlevent.yilmaz@pa.edu.tr

**Öz**

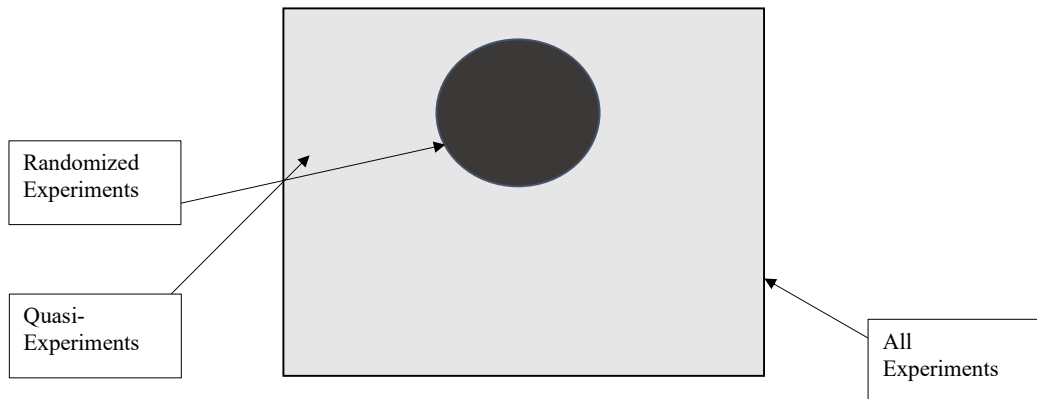
Eđitim ekonomisi alanı, eđitimin sürdürülebilir uzun vadeli ekonomik büyüme üretmesi ve insanların becerilerini artırması gerektiđini belirten büyüme teorilerinin ortaya çıkmasından sonra yoğun ilgi görmekte-dir (Hanushek ve Woessmann, 2015; Mincer, 1974). Eđitimsel arařtırmalar, evrensel önemine rađmen, sađlam nicel arařtırmaların nadiren kullanılması ve politika için istatistiksel olarak anlamlı nedensel çıkarımların eksikliđi nedeniyle tıp gibi alanlarla aynı ölçekte yerleřtirilmemiřtir (Creemers, Kyriakidēs ve Sammons, 2010). Bu göz önüne alındıđında, Amerika Birleřik Devletleri (Scimago Journal & Country Rank, 2018) gibi eđitim arařtırmalarında önde gelen ülkeler “tekrar tekrar bilimsel temelli arařtırmanın temeline dayanan eđitim politikası çağrıları” bařlatıp (Angrist, 2003, para 1) eđitim alanını nedensel çıkarımlar yapmak için titiz ve yenilikçi metodolojik yöntemler ve deneyler kullanmaya itmeye alıřmaktadır (Murnane ve Willett, 2011). Belirli bir arařtırma alanındaki ama test edilmek istenen müdahale-nin etkisini bulmaksa, deneysel bir yöntem kullanmak en uygundur (Beach ve Pedersen, 2016). Bu ned-enle, bu makale eđitim müdahalelerinde deneysel yöntemlerin, özellikle öđrenci sonuçlarını artırmak için finansal teřviklerin etkisini test eden randomize kontrol denemelerinin ve yarı deneysel tasarımların kul-lanılmasının fırsatlarını ve zorluklarını ele alacaktır. Birinci bölüm, deneysel tasarımlara genel bir bakıř, ardından randomize kontrol denemeleri ve yarı deneyleri tanımlayan ve bu yöntemleri kullanan ampirik alıřmaları tartıřan bölümlerdir. Her iki yöntem de titizlikle uygulandıđında kaliteli arařtırma üretebi-leceđinden, bu makalenin ne nicel ne de nitel arařtırma yöntemleri lehinde tartıřmadıđını belirtmek gere-kir (Lodico, Spaulding ve Voegtle, 2010).

**Anahtar Kelimeler:** Eđitim Ekonomisi; Arařtırma Yöntemleri; Deneysel Arařtırma Tasarımı; Eđitim So-nuçları; Rastgele Kontrollü alıřmalar; Yarı Deneyler; Mali Teřvikler

**JEL Codes:** I21, C80, C90

## 1. Overview of Experimental Design to Make Causal Inferences

As mentioned in Murnane and Willett (2011), 19th-century philosopher John Stuart Mill’s analy-sis state that the existence of a causal relationship is assumed “if (1) the cause preceded the effect, (2) the cause was related to the effect, and (3) we can find no plausible alternative explanation for the effect other than the cause” (Shadish, Cook, & Campbell, 2002, p.6). These philosophical noti- ons have carried on throughout centuries and entered into statistical literature as (1) ‘*Temporal Pre- cedence*’, (2) ‘*Covariation*’ and (3) ‘*No Alternative Explanation*’ (Cook & Campbell, 1979, p.37-38), and became validity measures for causal inference (Punch & Oancea, 2014). The most appropriate way to initiate research that can uphold all Mills’ three conditions is to run an experiment (Murnane & Willett, 2011), which is defined as a research design where “(a) one or more independent variables are manipulated to study their effect on a dependent variable and (b) participants are randomly as- signed to treatment or comparison groups” (Punch & Oancea, 2014, p. 378). To further elaborate, the independent variable is the intervention variable that is used to predict the dependent variable, which is the outcome that the research aims to measure (Thompson, 2006). Two basic types of expe- rimental research designs, as illustrated in Figure 1, are randomized and quasi-experiments (Mur- nane & Willett, 2011).

**Figure 1.** Two Kinds of Experiments

Source: (Murnane & Willett, 2011, p.30)

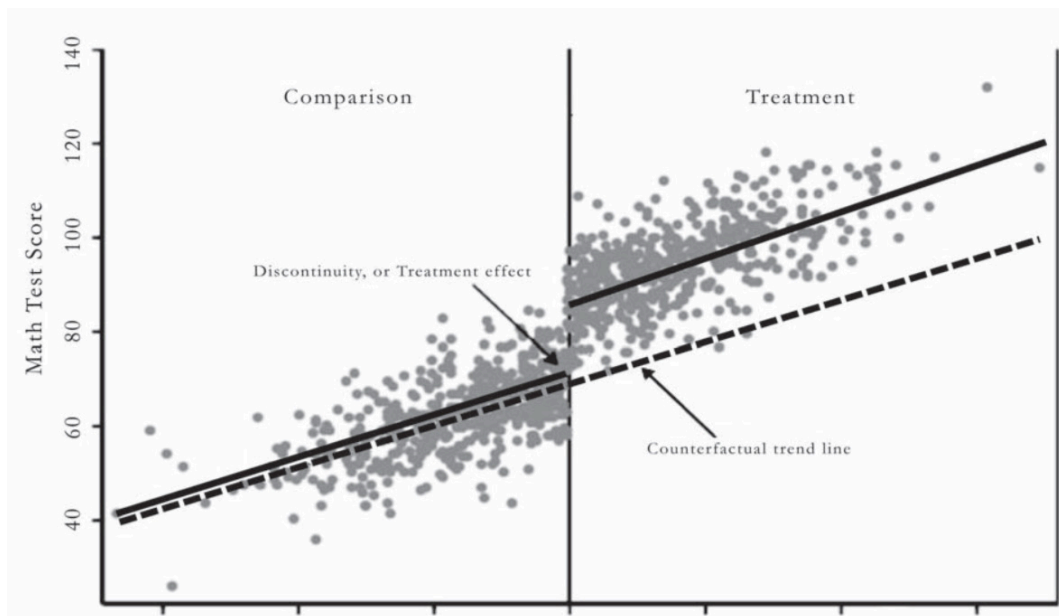
## 2. Randomized Control Trials

The randomized control trial (hereafter RCT) is regarded as the ‘gold standard’ for research dealing with measuring the impact of an intervention or prevention (Wes, et al., 2008), and the assessment of public policies (Cowen, Virk, Mascarenhas-Keyes, & Cartwright, 2017). While its theoretical foundation is analogous to any other experiment where participants are bifurcated into *intervention* (treatment) and *observation* (no treatment) groups (Kabish, Ruckes, Seibert-Grafe, Blettner, 2011), the reason RCTs are celebrated as the gold standard of research is that the assignment of participants into groups is entirely randomized (Kendall, 2003). The randomization process can take place from a simple coin toss to complex computational randomization algorithms, and the participants are unaware regarding which group they are assigned to (Shadish, Cook, & Campbell, 2002). It is at utmost importance for any researcher utilizing quantitative methods to establish a robust rigor within their statistical models by ensuring internal validity, using reliable data sets and a gathering a sample in which the findings are generalizable (Mujis, 2011). Statistical analysis is always prone to myriad problems such as multicollinearity where independent variables have high levels of correlation between one another (Gray, 2018; Pallant, 2016), and any finding where statistical framework misspecifications are not minimized and that the findings lack validity hinders the researcher from making causal inferences (Shaughnessy, Zechmeister, & Zechmeister, 2015). RCT, on the other hand, are designed in such a way to eliminate various validity problems and increase the confidence intervals of the output of the research (Punch & Oancea, 2014).

RCTs predominantly increase the statistical confidence of the research by minimizing a major internal validity threat of *selection bias* (Shadish, Cook, & Campbell, 2002). Given that the selection of participants into either treatment or no treatment groups is randomized, the risk of selection bias is

strongly minimized<sup>1</sup> (Every-Palmer & Howick, 2014). Furthermore, another important benefit of using RCTs to conduct quantitative research is the minimization of another prevent another risk of *confounding variables* (also referred to as ‘confounders’) (Attia, 2005). Confounders are variables “other than the one studied, that can cause or prevent the outcome of interest” (Attia, 2005, p. 259). In layman explanation, confounders can cause to research output to erroneously provide results that a change in a dependent variable is caused by an independent variable that in reality has no causal relationship to the dependent variable (Shadish, Cook, & Campbell, 2002). Exempli gratia, a research output could show a positive correlation between shark attacks and ice cream sales where the correct causal relationship would be with the increasing temperatures as it would both lead to an increase in people swimming in the ocean and purchasing ice cream (Martin, 2003). Hence, while confounders can be related to the outcome by being correlated with the dependent variable, that relationship is never causal (Cook & Campbell, 1979). What the randomization process of RCTs enable researchers is to distribute these confounding variables between both groups so that they cancel each other out by equally impacting both groups. Figure 2 demonstrates a hypothetical research output “results of a treatment designed to increase math test scores where (emphasis added) the discontinuity in the solid line indicates a treatment effect” (Gorrall, Curtis, Little, & Panko, 2015, p.22).

**Figure 2.** Hypothetical RCT Output Showing the Impact of the Treatment

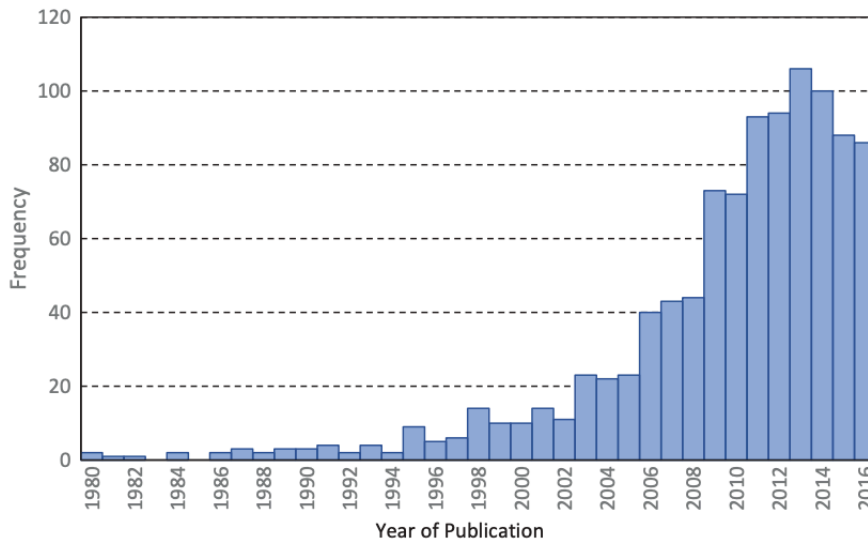


Source: (Gorrall, Curtis, Little, & Panko, 2015, p.22)

1 It is theoretically impossible to assert randomization entirely eliminates selection bias due to the possibility of poorly constructed randomization, or the recruiter’s/participant’s ability to guess the randomization (Kahan, Rehal, & Cro, 2015).

The usage of RCTs in response to the mandates disseminated by leading countries in educational research such as the United States (Scimago Journal & Country Rank, 2018), Australia and the UK (Besley, 2009) urging for scientific research in education has significantly increased, as illustrated in Figure 3, in the 21<sup>st</sup> century (Angrist, 2003; Murnane & Willett, 2011). In their systematic review, Connolly, Keenan, and Urbanska (2018) find that “a total of 1017 unique RCTs in the field of education have been (emphasis added) completed and reported between 1980 and 2016” (p.1). Majority of these studies have been constituted around the research questions of the impact of financial incentives/vouchers on student outcomes and various teaching methods (Angrist, 2003; Connolly, Keenan, & Urbanska, 2018). For instance, Angrist and Lavy (2002) conducted a randomized control trial in Israel where they identified 40 under-achieving high schools and then randomly selected 20 of them for treatment where students were eligible to receive \$1429 if they pass their high school matriculation exams called *Bagrut*. Their results show that 27% of the treatment group completed their Bagrut exams, in juxtaposition to the students in the control group where the exam completion rate was around 18% (Angrist & Lavy, 2002). Furthermore, students in the treatment group were around 8% more likely to complete the exam, which the authors state that the economic returns of such achievement for Israel is far beyond the matriculation exam awards (Angrist & Lavy, 2002; Angrist, 2003).

**Figure 3.** Number of RCTs in Education Completed Internationally Between 1980 and 2016

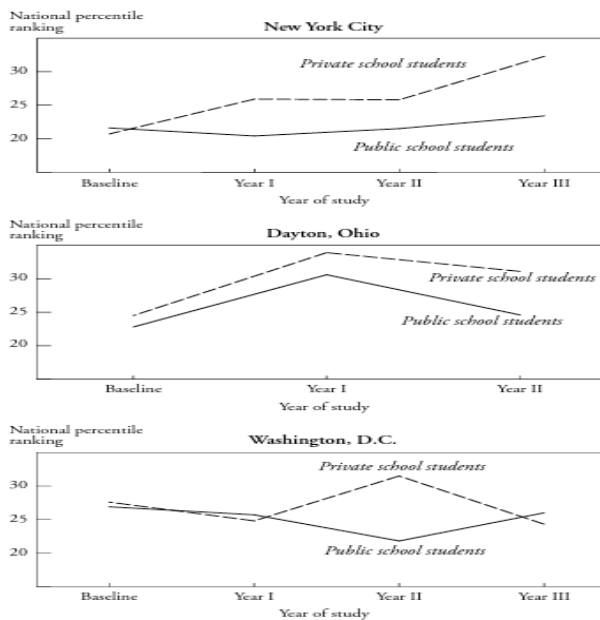


Source: (Connolly, Keenan, & Urbanska, 2018), p.10)

Another study conducted by Howell and Peterson (2006) focuses on the impact of school vouchers on the achievement of minority students in the United States and is claimed to be the most extensive randomized control trial in the field of education concerning vouchers (Howell & Peterson, 2018). As mentioned in Murnane and Willett (2011), the voucher theory gained prominence after

Nobel Prize Laureate economist Milton Friedman’s (1968) book called *Capitalism and Freedom*. Libertarian at the core, Friedman (1968) argues in favor of freedom of choice pervasively throughout his book. His statements in education reflects such libertarian approach as he argues that the US government should distribute equal value schooling vouchers to parents with children at the age of school (Friedman, 1968; Murnane & Willett, 2011). The choice of school selection, Friedman (1968) states, should be left to the parents as they would have the option to either subsidize all the expenses of their kids at a public school, or a portion of the costs at a private school. In their book called *The Education Gap*, Howell and Peterson (2006) test the impact of school vouchers and reflect on their study of more than 60,000 students, a majority of which are minorities coming from disadvantaged financial backgrounds in the cities of Dayton (Ohio), New York City, and Washington D.C. The authors find that the treatment showed a significant positive impact to African American students (hereafter AAS), as illustrated in Figure 4, while showing no statistical impact to other “ethnic groups, either Hispanics in New York City, or whites in Dayton” (p.166). The authors report that AAS in treatment in Dayton that switched to private schools scored around six percentage points higher on the “Iowa Tests of Basic Skills” (p.166) than the ones in control group who stayed at public schools. Furthermore, AAS in treatment in New York City and Washington D.C. that switched to private schools “scored “0.18 and 0.28 (emphasis added) standard deviation higher than the students in the control group” (p.166) in years one and two respectively. This study had a profound impact in the US education systems as states such as Wisconsin initiated voucher programs in the city of Milwaukee directed towards African American families (Howell & Peterson, 2006; 2018).

**Figure 4.** “African American Students’ Test Scores in Public and Private Schools in Three Cities”



Source: (Howell & Peterson, 2006, p.147)

RCTs face myriad of ethical and practical challenges and criticisms, the most commonly regarding the claim that randomization is unethical due to lack of knowledge regarding the efficacy of the treatment. (Hutchison & Styles, 2010). Hutchison and Styles (2010) argue that both educational and financial incentives can be offered to the participants since it is “rarely known what works [in educational research] in advance” (p.5). The authors also state that blinding, a key component of RCTs, is difficult to establish in “educational research trials due to the propensity for intervention group participants to ‘try harder’ (known as the Hawthorne Effect). However, other evaluation methods will rarely improve on this since they will also not operate blinding” (p.5). Further criticisms towards RCTs seem to be emanating from the staunch proponents of qualitative educational research. In their book of *Research Methods in Education*, Cohen, Manion, and Morrison, K. (2018) claim that it is impossible to use RCTs in educational research as cannot take place outside of isolated places due to inability to control outside factors, and further state that “randomized controlled trials belong to a discredited view of science as Positivism” (p. 314) (also mentioned in Connolly, Keenan, & Urbanska, 2018). The claim that RCTs are impossible to be conducted in educational research is redundant as there are 1017 successful RCT studies in the field of education in the past 15 years (Connolly, Keenan, & Urbanska, 2018). Besides, blatantly discrediting positivism which has led to many scientific breakthroughs with no firm evidence is unlikely to be taken into consideration in academia.

### 3. Quasi-Experimental Research Design

As “the prefix *quasi* means ‘resembling’” (Price, Jhangiani, & Chiang, 2015, para. 1), quasi-experiments (hereafter QE) are similar to RCTs given both are experimental methods, with the significant difference being that the participants are not randomized in QE design (Shadish, Cook, & Campbell, 2002). The researcher is autonomous regarding the assignment of participants to treatment or control groups by commonly using cutoffs or other criteria to balance both groups out (Punch & Oancea, 2014). Since the establishment and the operation of the randomization process is reasonably complex, QEs are more straight-forward to conduct; less costly and used more frequently in intervention research (Murnane & Willett, 2011). QEs usually have larger sample sizes given the absence of randomization<sup>2</sup> (Cook & Campbell, 1979), which allows the researcher to increase the generalizability of their findings, one of the main ingredients of rigor in quantitative research (Muji, 2011). While this may be true, the absence of randomization however exogenously causes QEs to face more internal validity problems in juxtaposition to RCTs (Punch & Oancea, 2014). Given the randomization process allows major internal validity problems such as confounding variables and selection bias to be minimized, these problems are pervasive in QEs (Thompson, 2006). The existence of confounding variables and selection bias, therefore, lowers the quality of the data (White & Sabarwal, 2014), hence the statistical significance of QEs in comparison to RCT, leading to lower chance to make causal inferences (Shadish, Cook, & Campbell, 2002).

A crucial aspect in research where QEs are far more favorable as opposed to RCTs is regarding ethical considerations (White & Sabarwal, 2014). Since no randomization takes place, researchers

---

2 As sample size increases, so does the complexity of conducting an RCT, given it would be harder to protect the randomization process for more people involved in research as participants (Gorrall, Curtis, Little & Panko, 2015).

usually take many aspects of participants' statuses and conditions into account such as giving them the access to choose between treatment or control, which lowers ambiguity regarding the expected outcomes on the participants and determining the participants who need the treatment the most (West, et al., 2008). Furthermore, as White and Sabarwal (2014) explains:

Quasi-experimental methods that involve the creation of a comparison group are most often used when it is not possible to randomize individuals or groups to treatment and control groups. This is always the case for ex-post impact evaluation designs. It may also be necessary to use quasi-experimental designs for ex-ante impact evaluations, for example, where ethical, political or logistical constraints, like the need for a phased geographical roll-out, rule out randomization (p.2).

Henceforward, the researcher can minimize possible ethical issues using QEs (Murnane & Willett, 2011). Furthermore, unlike RCTs, QEs can take place during or after the treatment has taken place, which can minimize another ethical issue that RCTs face regarding not knowing the effects and the outcome of the treatment (White & Sabarwal, 2014).

Given that QEs are more accessible to implement and also useful where RCTs cannot take place, they are pervasively adopted to study educational interventions (Murnane & Willett, 2011). Angrist, Bettinger, Bloom, King, and Kremer (2001) famously conducted one of the most substantial school voucher experimental studies in the world (Angrist, 2003) using data from more than 125,000 Colombian pupils coming from disadvantageous financial backgrounds. Although the vouchers were distributed using a lottery system as the authors state that the demand for the vouchers was far beyond supply (Angrist, Bettinger, Bloom, King, & Kremer, 2001), this does not qualify the study as an RCT given that the participants were aware of their group assignment. Although the authors do not provide the monetary value of these vouchers, they state that the "vouchers cover (emphasis added) somewhat more than half the cost of private secondary school" (Angrist, Bettinger, Bloom, King, & Kremer, 2001, p. 1535). A survey conducted three years after the voucher distribution show that while there were no major differences in the enrollment rates between the winners and the losers of the vouchers, "lottery winners were 15 percentage points more likely to attend private schools rather than public schools" (p.1536). Besides, the winners "completed an additional 0.1 years of school" (p. 1536) and were about ten percentage points to avoid great repetition and complete eighth grade. Finally, the winner group also scored 0.2 standard deviations higher than losers on standardized exams, and the effect of the vouchers were larger on girls and "more precisely estimated than the effect on boys" (p. 1536).

## Conclusion

This paper inquired the usage of experimental research methods of RCTs and a QEs in the field of education concerning educational interventions, mostly regarding financial incentives and vouchers to increase student outcomes. Section 1 gave an overview of experimental designs to make causal inference, since it is essential to understand the theoretical underpinnings of this type of research before delving into the methods. Following sections elaborated on the opportunities and challenges



of RCTs and QEs, delineated on the establishment and the implementation of these methods from a statistical standpoint. Various studies that employ these methods were discussed.

While both are experimental methods in their essence, the main trade-off between RCTs and QEs seems to be statistical significance vs. ethics. RCT's are the gold standard of research to test interventions (West et al., 2008), and public policy assessments (Cowen, Virk, Mascarenhas-Keyes, & Cartwright, 2017). They are likely to provide high levels of statistical significance and minimized validity threats such as selection bias and confounders (Punch & Oancea, 2014). Consequently; however, they are prone to many ethical dilemmas (Hutchison & Styles, 2010). QE, on the other hand, could be specified as the inverse function of RCTs as they are likely to come with lower levels of statistical significance and ethical issues, and higher levels of validity threats (White & Sabarwal, 2014). Overall, the profound impact of experimental methods on the prominence and credibility of the field of education, and on policy-making is undeniable. It is clear that governments ask for scientifically based research to implement policy changes (Angrist, 2003) as the quest for causal inferences is global (Murnane & Willett, 2011).

## BIBLIOGRAPHY

- ANGRIST, J. (2003, Summer). Randomized Trials and Quasi-Experiments in Education Research. Retrieved March 11, 2019, from <https://www.nber.org/reporter/summer03/angrist.html>
- ANGRIST, J. D. (2004). American Education Research Changes Tack. *Oxford Review of Economic Policy*, 20(2), 198-212. doi:10.1093/oxrep/grh011
- ANGRIST, J., BETTINGER, E., BLOOM, E., KING, E., & KREMER, M. (2001). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *THE AMERICAN ECONOMIC REVIEW*, 1535-1558. doi:10.3386/w8343
- ANGRIST, J., & Lavy, V. (2002). The Effect of High School Matriculation Awards: Evidence from Randomized Trials. *American Economic Review*. doi:10.3386/w9389
- ATTIA, A. (2005). Bias in RCTs: Confounders, selection bias and allocation concealment. *Middle East Fertility Society Journal*, 10(3), 258-261. Retrieved March 14, 2019.
- BEACH, D., & PEDERSEN, R. B. (2016). *Causal case study methods foundations and guidelines for comparing, matching and tracing*. Ann Arbor: University of Michigan Press.
- BESLEY, T. (2009). *Assessing the quality of educational research in higher education: International perspectives*. Rotterdam, Netherlands: Sense.
- BRADSHAW, C. P., ZMUDA, J. H., KELLAM, S. G., & Ialongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology*, 101(4), 926-937. doi:10.1037/a0016586
- COHEN, L., MANION, L., & MORRISON, K. (2018). *Research methods in education*. New York: Routledge.
- CONNOLLY, P., BIGGART, A., MILLER, S., OHARE, L., & THURSTON, A. (2017). *Using randomised controlled trials in education*. Thousand Oaks, CA: SAGE.
- CONNOLLY, P., KEENAN, C., & URBANSKA, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, 60(3), 276-291. doi:10.1080/00131.881.2018.1493353
- COOK, T. D., & CAMPBELL, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.

- COWEN, N., VIRK, B., MASCARENHAS-KEYES, S., & CARTWRIGHT, N. (2017). Randomized Controlled Trials: How Can We Know “What Works”? *Critical Review*, 29(3), 265-292. doi:10.1080/08913.811.2017.1395223
- CREEMERS, B. P., KYRIAKIDĒS, L., & SAMMONS, P. (2010). *Methodological advances in educational effectiveness research*. Abingdon, Oxon, England: Routledge.
- ECCLES, M., GRIMSHAW, J., CAMPBELL, M., & RAMSAY, C. (2003). Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care*, 12(1), 47-52. doi:10.1136/qhc.12.1.47
- EVERY-PALMER, S., & HOWICK, J. (2014). How evidence-based medicine is failing due to biased trials and selective publication. *Journal of Evaluation in Clinical Practice*, 20(6), 908-914. doi:10.1111/jep.12147
- FRIEDMAN, M. (1968). *Capitalism and freedom*. Chicago, Ill: University of Chicago Press.
- GORRALL, B., CURTIS, J., LITTLE, T., & PANKO, P. (2015). Alternatives to Randomized Control Trials: A Review of Three Quasi-experimental Designs for Causal Inference. *Actualidades En Psicología*, 29(119), 19-27. doi:10.15517/ap.v29i119.18810
- GRAY, D. E. (2018). *Doing research in the real world*. London: SAGE Publications.
- HANUSHEK, E. A., & WOESSMANN, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. Cambridge: The MIT Press.
- HOWELL, W. G., & PETERSON, P. E. (2006). *The education gap: Vouchers and urban schools*. Washington (D.C.): Brookings Institution Press.
- HOWELL, W. G., & PETERSON, P. E. (2018, SEPTEMBER 10). The Education Gap. Retrieved March 14, 2019, From <https://www.brookings.edu/book/the-education-gap/>
- HUTCHISON, D., & STYLES, B. (2010). *A Guide To Running Randomised Controlled Trials For Educational Researchers*. Slough: NFER.
- KABISCH, M., RUCKES, C., SEIBERT-GRAFE, M., & BLETTNER, M. (2011). Randomized controlled trials: part 17 of a series on evaluation of scientific publications. *Deutsches Arzteblatt international*, 108(39), 663-8.
- KAHAN, B. C., REHAL, S., & CRO, S. (2015). Risk of selection bias in randomised trials. *Trials*, 16(1). doi:10.1186/s13063.015.0920-x
- KENDALL, J. M. (2003). Designing a research project: Randomised controlled trials and their principles. *Emergency Medicine Journal*, 20(2), 164-168. doi:10.1136/emj.20.2.164
- LODICO, M. G., SPAULDING, D. T., & VOEGTLE, K. H. (2010). *Methods in Educational Research: From Theory to Practice*. San Francisco: Jossey-Bass.
- MARTIN, G. H. (2013, November 10). Randomized Control Trials and Confounding. Retrieved March 14, 2019, from <https://www.youtube.com/watch?v=7ybuE39BpQ8>
- MINCER, J. (1974). *Schooling, Experience And Earnings*. New York: National Bureau Of Economic Research.
- MUIJS, D. (2011). *Doing quantitative research in education with SPSS*. Los Angeles: Sage Publications.
- MURNANE, R. J., & WILLETT, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. Oxford: Oxford University Press.
- PALLANT, J. (2016). *SPSS survival manual: A step by step guide to data analysis using SPSS*. Maidenhead: Open University Press.
- PETERSON, P. E., & HOWELL, W. G. (2004). Efficiency, Bias, and Classification Schemes: A Response to Alan B. Krueger and Pei Zhu. *American Behavioral Scientist*, 47(5), 699-717. doi:10.1177/000.276.4203260158
- PRICE, P. C., JHANGIANI, R. S., & CHIANG, A. (2015, October 13). Quasi-Experimental Research. Retrieved March 15, 2019, from <https://opentextbc.ca/researchmethods/chapter/quasi-experimental-research/>

- PUNCH, K., & OANCEA, A. (2014). *Introduction to research methods in education*. London: SAGE Publications.
- SCIMAGO JOURNAL & COUNTRY RANK IN EDUCATION. (n.d.). Retrieved March 11, 2019, from <https://www.scimagojr.com/countryrank.php?category=3304>
- SHADISH, W. R., COOK, T. D., & CAMPBELL, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.
- SHAUGHNESSY, J. J., ZECHMEISTER, E. B., & ZECHMEISTER, J. S. (2015). *Research methods in psychology*. Dubuque: McGraw-Hill Education.
- THOMPSON, C. (2006). Research Study Designs: Experimental and Quasi-Experimental. *Air Medical Journal*, 25(6), 242-246. doi:10.1016/j.amj.2006.09/001
- VAUS, D. A. (2001). Chapter 4, Experimental Design. In *Research Design in Social Research* (pp. 53-69). Los Angeles, CA: Sage.
- WEST, ET AL. (2008). Alternatives to the randomized controlled trial. *American journal of public health*, 98(8), 1359-66.
- WHITE, H., & S. SABARWAL (2014). Quasi-experimental Design and Methods, Methodological Briefs: Impact Evaluation 8, UNICEF Office of Research, Florence.
- WOESSMANN, L., & PETERSON, P. E. (2014). *Schools and the Equal Opportunity Problem*. Cambridge: MIT Press.