

# Kausalität feststellen: Ein Plädoyer für randomisierte kontrollierte Studien

**Adrien Bouguen** [J-PAL Europe, PSE – abouguen@povertyactionlab.org]

**Marc Gurgand** [J-PAL Europe, Paris School of Economics – gurgand@pse.ens.fr]

Randomisierte kontrollierte Studien (englisch: *randomized controlled trials*, RCT) stellen eine experimentelle Methode dar, um die Auswirkungen von Bildungspolitik exakt zu bewerten. Die Erfahrung zeigt, dass RCTs in Europa bereits durchgeführt werden und dass Ergebnisse aus einem europäischen Land wiederum für andere Länder relevant sein können. Die Europäische Union könnte deshalb nach Möglichkeiten suchen, Experimente durchzuführen und deren Ergebnisse zu bewerten und zu verbreiten.

[Übersetzt aus dem Englischen von Ruth Maria Schueler.]

Da die Existenz langfristiger Effekte von Bildung auf verschiedene Aspekte des Erfolgs im Leben zweifelsfrei anerkannt wird, hat sich die Europäische Union seit langem dazu verpflichtet, im Bereich Bildung ehrgeizige Ziele umzusetzen. Angesichts der gegenwärtigen Krise der öffentlichen Ausgaben und dem bereits relativ hohen Anteil des BIPs für Bildungshaushalte in Europa, stellt sich jedoch die Frage, wie es den Mitgliedstaaten gelingen kann, Bildung im Allgemeinen zu verbessern ohne das Staatsdefizit übermäßig zu erhöhen.

Eine naheliegende Antwort wäre, umsichtig die effizienteste Politikmaßnahme zu wählen. Um Bildungsprogramme wie die Reduzierung der Klassengröße, Etablierung von Internaten, rehabilitierende Bildungsprogramme oder neue pädagogische Praktiken zu bewerten, braucht es allerdings präzise Maßnahmen und eine strikte Methodik. Derartige Programme sind üblicherweise mit hohen Kosten verbunden und einmal ausgeweitet, sind sie meist nur schwer wieder einzustellen. Eine gründliche Bewertung ihrer Wirksamkeit im Voraus wäre somit erstrebenswert.

## EINE TRANSPARENTE METHODOLOGIE

Der einfache Vergleich einer Gruppe von Bezugsberechtigten eines Bildungsprogramm, mit einer Gruppe Nicht-Bezugsberechtigter führt leider nicht zu brauchbaren Ergebnissen, da sich die beiden Gruppen grundsätzlich unterscheiden. Einige nicht-experimentelle Methoden (*Matching* oder multivariate Regression) sind zwar nützlich, berücksichtigen allerdings kaum unbeobachtete Eigenschaften, die sich zwischen der Gruppe der Bezugsberechtigten und der Nicht-Bezugsberechtigten unterscheiden können (wie z.B. Motivation, Intelligenz, Einsatzbereitschaft, ...). Es ist jedoch entscheidend, solche unbeo-

bachteten Eigenschaften im Bildungskontext zu berücksichtigen.

Ein alternativer Weg, um die Auswirkungen von Bildungsprogrammen zu evaluieren (z.B. die Reduzierung der Klassengröße oder rehabilitierende Bildungsprogramme, ...) ist es, ein experimentelles Setting aufzubauen – eine sogenannte randomisierte kontrollierte Studie (RCT) – bei der die Ergebnisgrößen (Testergebnisse, Abbruchraten, usw.) der vom Programm begünstigten Schüler (die sogenannte „Behandlungsgruppe“) mit dem Ergebnis der in ihren Eigenschaften ähnlichen Nicht-Begünstigten (die sogenannte „Kontrollgruppe“) verglichen werden. Um sicherzustellen, dass beide Gruppen schon vor Programmstart statistisch identisch sind, werden die Gruppen der teilnahmeberechtigten Schüler (oder Schulen) zufällig festgelegt.

## AKTUELLE EUROPÄISCHE BEISPIELE

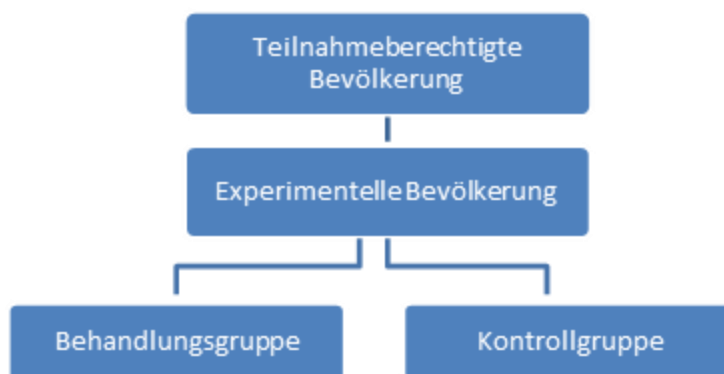
In den letzten Jahren wurden im Bildungsbereich verschiedene groß angelegte Experimente in europäischen Ländern durchgeführt. Beispielsweise evaluierte die französische Regierung 2007 ein Programm für die Einbindung von Eltern von Sechstklässlern, das für eine zufällig ausgewählte Untergruppe von 200 Klassen umgesetzt wurde. Man konnte zeigen, dass dieses einfache Programm das Verhalten von Schülern stark verbessert hat, vor allem in Bezug auf Schulversäumnisse. Im Zuge dieser positiven Ergebnisse entschied sich die französische Regierung, das Programm auszuweiten.

Seit September 2012 bewertet die dänische Regierung den Einfluss von zusätzlichen Lehrern, die reguläre Lehrer in den Unterrichtsfächern Dänisch und Mathematik unterstützen sollen. Von 105 Schulen, werden 35 Schulen einen zusätzlichen Lehrer erhalten, 35 Schulen werden eine zusätzliche Person anstellen, die nicht zwangsweise ein Lehrer sein

### Externe und interne Gültigkeit von randomisierten kontrollierten Studien (RCTs)

Aus der teilnahmeberechtigten Bevölkerung wird eine Untergruppe von Individuen oder Schulen ausgewählt.

Die zufällige Zuteilung stellt sicher, dass beide experimentellen Gruppen die gleichen Eigenschaften aufweisen.



muss, und 35 Schulen werden als Kontrollgruppe dienen. Diese Alternative ist kostengünstiger als die Klassengröße zu reduzieren. Ihr Einfluss ist bisher jedoch noch nicht bekannt.

#### VIELE MÖGLICHKEITEN ZUR RANDOMISIERUNG

Obwohl RCTs in vielen Bereichen Anwendung finden, eignet sich die Methode besonders gut um neue bildungspolitische Maßnahmen zu bewerten. Tatsächlich können RCTs im Gegensatz zu nicht-experimentellen Methoden, welche Politikmaßnahmen auch bewerten können nachdem sie bereits gestartet sind, nur für neue Bildungsprogramme verwendet werden, die noch nicht großflächig ausgeweitet wurden. Zudem setzen RCTs voraus, dass die Gruppe von Teilnehmern und Nicht-Teilnehmern vor Programmstart zufällig festgelegt wird.

Diese zweite Beschränkung wird häufig als ethisch und politisch bedenklich erachtet. Im Prinzip bedeutet der Ausschluss keine Nutzeneinbußen, da die Auswirkungen des Programms noch nicht bekannt sind. In vielen Fällen könnten die Kandidaten jedoch die Teilnahme an dem Programm per se wertschätzen, ohne den möglichen Einfluss des Programms überhaupt zu kennen (man denke nur an ein Geldtransferprogramm). Einige Gelegenheiten ermöglichen jedoch eine zufällige Zuteilung. Wird ein Programm vor einer Ausweitung beispielsweise zunächst nur in wenigen Schulen eingeführt, ist es relativ zweckdienlich, zufällig zuzuteilen, welche Schulen zuerst von dem Programm profitieren (*phase-in* Methode). Gleichermassen scheint es natürlich, zufällig eine Anzahl an Schülern aus einem Pool von teilnahmeberechtigten Kandidaten auszuwählen, wenn die Nachfrage nach den Dienstleistungen aus dem Programm das Angebot übersteigt (z.B. bei einer

begrenzten Anzahl an Schlafplätzen in einem Internat). Eine zufällige Zuteilung wird häufig akzeptiert und sogar als fairer betrachtet als eine Auswahl an Hand von *ad hoc* Kriterien.

#### EXTERNE GÜLTIGKEIT UND UMSETZUNG IN DER EUROPÄISCHEN UNION

Wenn RCTs angemessen konzipiert werden, generieren sie zuverlässige, robuste und präzise Schätzer, die den Einfluss des Programms messen können: In anderen Worten bezeichnet man sie als „intern gültig oder valide“ (s. Abbildung). Da RCTs jedoch lokal implementiert werden, kann es manchen Evaluierungen an „externer Gültigkeit“ fehlen und es ist daher nicht immer klar, ob sich das Ergebnis einer spezifischen Evaluierung auf andere ökonomische, soziale und institutionelle Kontexte übertragen lässt. Für dieses Problem, das in keiner Weise spezifisch für RCTs ist, gibt es keine einfache Lösung.

Ein möglicher Weg, die Ergebnisse eines spezifischen Kontexts zu verifizieren, ist es, das gleiche Programm an anderer Stelle zu replizieren und zu evaluieren. Mit Blick auf die offene Koordinierungsmethode auf europäischer Ebene, würde die Europäische Union einen perfekten Schauplatz für eine Wiederholung von Programmen in anderen Ländern bieten, welche in einem Mitgliedstaat bereits als erfolgreich bewertet werden konnten. Europäische Länder sehen sich vielen ähnlichen Bildungsthemen gegenüber und könnten davon profitieren, bei Experimenten zu kooperieren, Ergebnisse zu verbreiten und für effiziente Politikmaßnahmen zu werben.

Für weitere Details vgl.: Adrien Bouguen, Marc Gurgand, *Randomized Controlled Experiments in Education*. EENEE Analytischer Bericht Nr. 11, Februar 2012, [http://www.eenee.de/dms/EENEE/Analytical\\_Reports/EENEE\\_AR11.pdf](http://www.eenee.de/dms/EENEE/Analytical_Reports/EENEE_AR11.pdf).

Europäisches Expertennetzwerk Bildungsökonomik (EENEE)

Gefördert durch die Europäische Kommission, Generaldirektion Bildung und Kultur



EENEE-Veröffentlichungen und weitere Informationen unter: <http://www.eenee.org>. – Koordination: Ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München e.V., Poschingerstr. 5, 81679 München, Tel.: +49-89-9224-1388, E-mail: [eenee@ifo.de](mailto:eenee@ifo.de). © EENEE 2012. Alle Rechte vorbehalten. Die hier vertretenen Meinungen sind die der Verfasser und nicht der Europäischen Kommission.