

## Evaluer la Causalité : les Atouts de l'Essai Randomisé Contrôlé

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

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

L'Essai Randomisé Contrôlé (ERC) est une méthode expérimentale utilisée pour évaluer rigoureusement l'effet des politiques d'éducation. L'expérience montre que la pratique des ERC se développe en Europe et que les résultats pour un pays peuvent s'appliquer à d'autres. L'Union Européenne pourrait donc saisir les opportunités que représentent l'expérimentation, l'évaluation et la diffusion des résultats.

[Traduit de l'anglais par Alexis Maitre.]

L'Union Européenne s'est depuis longtemps fixée d'ambitieux objectifs en matière d'éducation, tant les bénéfices de long terme de cette dernière sont fermement établis. Cependant, du fait de la crise de la dépense publique et de la part déjà importante des budgets européens consacrée à l'éducation, comment les Etats membres peuvent-ils améliorer le niveau d'éducation général sans augmenter le déficit public?

Pour ce faire, il suffit à première vue de sélectionner les politiques les plus efficaces. Toutefois, des mesures précises et des méthodes rigoureuses doivent être appliquées pour évaluer l'impact de mesures telles que la réduction de la taille des classes, de l'internat, des cours de rattrapage ou encore des nouvelles méthodes pédagogiques.

### UNE METHODOLOGIE TRANSPARENTE

Comparer les résultats d'un groupe de bénéficiaires d'une mesure donnée avec un groupe de non-bénéficiaires ne permet malheureusement pas d'obtenir de résultat utile, les deux groupes tendant à être intrinsèquement différents. Des méthodes non-expérimentales (appariement ou régression multivariée) existent pour répondre à ce problème, mais elles présentent l'inconvénient de ne pas permettre la prise en compte de caractéristiques inobservables (motivation, intelligence, dynamisme, ...) qui peuvent être inégalement réparties entre les bénéficiaires et les non-bénéficiaires. De telles caractéristiques sont pourtant centrales dans le contexte de l'éducation.

Une alternative pour évaluer l'effet d'une politique d'éducation consiste à établir un cadre expérimental – un Essai Randomisé Contrôlé (ERC) – où les résultats (score au test, taux d'échec, ...) des élèves bénéficiaires du programme (appelés « groupe de traitement ») sont comparés à ceux d'élèves similaires mais ne bénéficiant pas de la mesure (appelés « groupe de contrôle »). Pour s'assurer

que les deux groupes sont ex-ante statistiquement identiques, l'éligibilité au traitement pour un élève (ou une école) est déterminée aléatoirement.

En effet, si les groupes sont déterminés aléatoirement, on peut penser qu'ils sont statistiquement identiques avant la mise en place de la mesure. Toute différence de résultats entre les deux groupes à la fin de l'expérience peut donc être attribuée à la mesure : c'est l'effet du traitement.

### RECENTS EXEMPLES EN EUROPE

Ces dernières années, plusieurs expériences à grande échelle ont été effectuées dans des pays européens. Par exemple, en 2007, le gouvernement français a évalué un programme d'implication des parents pour des élèves de 6<sup>ème</sup> auquel 200 classes choisies aléatoirement ont participé. Il a été prouvé que cette simple mesure améliore l'attitude des élèves, en particulier en ce qui concerne l'absentéisme. A la vue de ces résultats encourageants, le gouvernement français a décidé de généraliser la mesure.

Depuis septembre 2012, le gouvernement danois évalue quant à lui l'impact de la présence d'enseignants supplémentaires pour aider les écoles à gérer l'apprentissage des mathématiques et du danois. Sur un échantillon de 105 écoles, 35 se sont vues affecter un nouveau professeur, 35 ont embauché une personne ressource supplémentaire qui n'était pas forcément un professeur et 35 écoles constituent le groupe de contrôle. Il s'agit d'une mesure certes moins coûteuse que la réduction de la taille des classes mais dont l'efficacité reste à prouver.

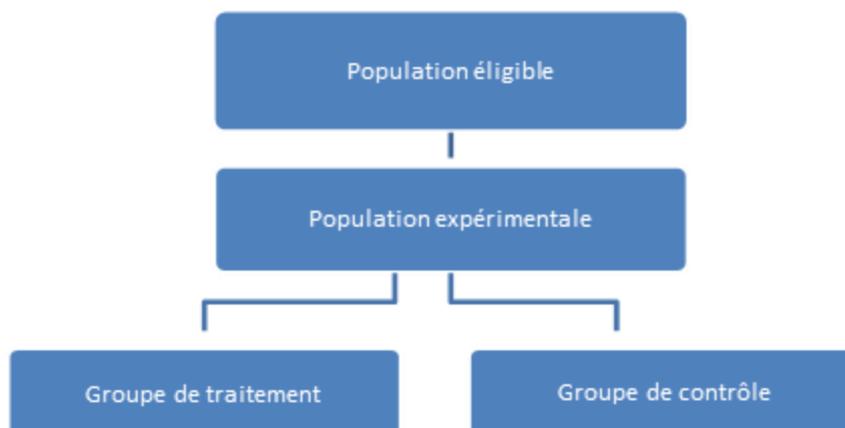
### DE NOMBREUSES OPPORTUNITES POUR RANDOMISER

Bien que les ERC soient utilisés dans de nombreux domaines, cette méthode est particulièrement adaptée à l'évaluation de nouvelles politiques en matière d'éducation.

### Validité externe et interne des Essais Randomisés Contrôlés (ERC)

Un échantillon d'élèves ou d'écoles est tiré parmi la population éligible.

L'affectation aléatoire au traitement garantit que les deux groupes ont les mêmes caractéristiques.



En effet, à l'inverse des méthodes non-expérimentales qui ne peuvent évaluer les mesures qu'après leur mise en place, l'ERC peut seulement être utilisé sur programmes qui n'ont pas encore été généralisés. De plus, l'usage des ERC requiert que les groupes de traitement et de contrôle soient formés aléatoirement.

Cette dernière contrainte est souvent considérée comme éthiquement ou politiquement délicate. En principe, exclusion des candidats par affectation aléatoire au programme ne les prive d'aucun bénéfice puisque les effets du programme sont par essence inconnus. Il reste que certains candidats accordent une valeur à la participation en elle-même sans considération d'effet du traitement (une prime d'allocations, par exemple). Il existe tout de même de nombreuses opportunités pour effectuer cette allocation aléatoire. Par exemple, lorsqu'un programme est introduit dans quelques écoles avant d'être généralisé, il est relativement aisé de sélectionner aléatoirement les écoles participantes (méthode d'introduction progressive). De la même manière, quand la demande pour le service fourni par le programme excède l'offre (nombre restreint de lits dans un internat par exemple), il est plutôt naturel de sélectionner aléatoirement un certain nombre d'élèves parmi la liste des candidats éligibles. L'affectation aléatoire au traitement est le plus souvent très bien acceptée et peut

même être considérée plus juste que la sélection sur la base de critères *ad hoc*.

#### VALIDATION EXTERNE ET REPRODUCTION A L'ECHELON EUROPEEN

Lorsqu'ils sont conçus correctement, les ERC produisent des estimations fiables, robustes et précises des effets d'une mesure : en d'autres termes, ils jouissent d'une « validité interne » (voir la figure). Cependant, puisque les ERC sont effectués localement, certains résultats peuvent manquer de « validité externe ». Pour le dire autrement, il n'est pas certain qu'un résultat spécifique puisse s'appliquer à d'autres contextes institutionnels, économiques et sociaux. Ce problème, qui n'est en aucun cas spécifique aux ERC, n'a pas de solution simple.

Une voie possible pour s'en défaire consiste à reproduire et évaluer le même programme dans différents contextes. En accord avec la méthode ouverte de coordination, l'Union Européenne pourrait ainsi être le parfait acteur pour reproduire dans d'autres Etats membres un programme ayant réussi dans un Etat. Les pays européens font en effet face à des défis communs en matière d'éducation et pourraient tirer profit de la coopération autour des expériences, de la diffusion des résultats et de la quête pour des politiques publiques plus efficaces.

Pour plus de détails, voir: Adrien Bouguen, Marc Gurgand, *Randomized Controlled Experiments in Education*. Rapport Analytique No. 11 de l'EENEE, Février 2012, [http://www.eenee.de/dms/EENEE/Analytical\\_Reports/EENEE\\_AR11.pdf](http://www.eenee.de/dms/EENEE/Analytical_Reports/EENEE_AR11.pdf).