

Effets du Réseau d'enseignement prioritaire à Genève : comment mesurer les effets d'un dispositif implanté en plusieurs phases ?

Anne Soussi et Gianreto Pini

Volume 40, numéro 1, 2017

Réception : 31 août 2015

Version finale : 23 janvier 2017

Acceptation : 7 février 2017

URI : <https://id.erudit.org/iderudit/1041005ar>

DOI : <https://doi.org/10.7202/1041005ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Soussi, A. & Pini, G. (2017). Effets du Réseau d'enseignement prioritaire à Genève : comment mesurer les effets d'un dispositif implanté en plusieurs phases ? *Mesure et évaluation en éducation*, 40(1), 101–128.

<https://doi.org/10.7202/1041005ar>

Résumé de l'article

Les précédentes évaluations n'ont pas permis de mettre en évidence des effets de la politique d'éducation prioritaire implantée à Genève (Réseau d'enseignement prioritaire) sur les compétences des élèves, pas plus en les comparant avec celles d'élèves scolarisés dans d'autres établissements ni même d'établissements présentant des caractéristiques sociodémographiques similaires n'ayant pas bénéficié du dispositif. Dans cet article, nous allons explorer une voie inspirée de l'analyse des séries temporelles afin de mesurer les effets d'un dispositif d'éducation prioritaire mis en place de manière progressive et de tenir ainsi compte du temps d'exposition différent des élèves. La méthode utilisée consiste à considérer pour chaque établissement quatre moments différents. Ce plan expérimental permet de répondre aux questions suivantes : Existe-t-il une évolution positive des résultats des élèves après l'introduction du dispositif ? Qu'en est-il deux ans plus tard ?

Effets du Réseau d'enseignement prioritaire à Genève: comment mesurer les effets d'un dispositif implanté en plusieurs phases?

Anne Soussi

*Service de la recherche en éducation, Département de l'instruction
publique, de la culture et du sport, Genève*

Gianreto Pini

Université de Genève

MOTS CLÉS: séries temporelles, éducation prioritaire, compétences des élèves, école primaire

Les précédentes évaluations n'ont pas permis de mettre en évidence des effets de la politique d'éducation prioritaire implantée à Genève (Réseau d'enseignement prioritaire) sur les compétences des élèves, pas plus en les comparant avec celles d'élèves scolarisés dans d'autres établissements ni même d'établissements présentant des caractéristiques sociodémographiques similaires n'ayant pas bénéficié du dispositif. Dans cet article, nous allons explorer une voie inspirée de l'analyse des séries temporelles afin de mesurer les effets d'un dispositif d'éducation prioritaire mis en place de manière progressive et de tenir ainsi compte du temps d'exposition différent des élèves. La méthode utilisée consiste à considérer pour chaque établissement quatre moments différents. Ce plan expérimental permet de répondre aux questions suivantes: Existe-t-il une évolution positive des résultats des élèves après l'introduction du dispositif? Qu'en est-il deux ans plus tard?

KEY WORDS: time series analysis, priority education, students' skills, primary school

Previous research did not highlight any positive effects of Geneva Educational Priority Policy on students' skills (the Réseau d'enseignement prioritaire). Students in educational priority school underperformed compared to their counterparts in other schools, even when these schools had similar sociodemographic characteristics. In this article, we will explore a different method to measure the effects of this progressive educational priority device. More specifically, we will use a method based on time series analysis, and thus take account of different exposure time of students. This method consists in four different times for each school. This experimental design will give answers to the following questions: Is there a positive evolution of students' results after the introduction of the device? And two years later?

PALAVRAS-CHAVE: séries temporais, educação prioritária, competências dos alunos, escola primária

As avaliações anteriores não conseguiram demonstrar os efeitos da política de educação prioritária implementada em Genebra (criação da Rede de ensino prioritário) sobre as competências dos alunos, nem compará-los com os efeitos nos alunos de outras escolas nem mesmo com escolas que apresentam características sociodemográficas similares. Neste artigo, vamos explorar um caminho inspirado pela análise de séries temporais para medir os efeitos do dispositivo de educação prioritária implementado gradualmente e, assim, ter em conta o diferente tempo de exposição dos alunos. O método utilizado consiste em considerar para cada estabelecimento quatro momentos diferentes. Este plano experimental permite responder às seguintes questões: Há uma evolução positiva dos resultados dos alunos após a introdução do dispositivo? E dois anos mais tarde?

Introduction

La plupart des pays européens et d'Amérique du Nord ont mis en place des politiques d'éducation prioritaire pour lutter contre les inégalités, et ce, depuis les années 1960 ou 1970. En règle générale, leurs effets positifs sur les acquis des élèves n'ont pu être montrés de façon évidente (Demeuse, Frandji, Greger & Rochex, 2008, 2011 ; Dutrévis & Crahay, 2009). Cela explique en partie le fait que ces politiques ont évolué au cours des décennies (Rochex, 2008). Le Réseau d'enseignement prioritaire (REP) mis en place à Genève depuis 2006 ne fait pas exception. Il s'est constitué progressivement et comporte, depuis 2011, 17 établissements primaires sur 81, ce qui représente environ 20% des élèves scolarisés à l'école primaire. Les conditions d'entrée d'un établissement dans le REP sont principalement liées à la proportion d'élèves provenant de milieux défavorisés, qui doit être d'au moins 55%. Une seconde condition est utilisée en cas de situation « limite » (c.-à-d. une proportion d'enfants d'ouvriers et divers/sans indication se situant entre 51 et 54%) : la proportion d'élèves allophones (plus de 55% d'allophones scolarisés dans l'établissement).

Les principales mesures du REP sont les suivantes : un meilleur encadrement des élèves (diminution des effectifs par classe à raison de deux élèves, plus grande dotation en enseignants chargés de soutien pédagogique) et, surtout, la création de la fonction d'éducateurs scolaires permettant de développer les liens entre famille et école. À côté de ces mesures spécifiques, l'enseignement primaire a connu une réforme globale avec l'instauration de directions, de projets et de conseils d'établissement. Ces différents éléments ont d'abord été expérimentés dans le REP, puis généralisés à l'ensemble des établissements primaires.

Les différentes études menées depuis son introduction (Soussi, Guignard, Hayoz & Nidegger, 2008 ; Soussi & Nidegger, 2010, 2015) n'ont pas permis de démontrer des effets positifs marqués concernant les acquis scolaires. Il faut d'emblée souligner que la mesure des effets de ces politiques de discrimination positive pose des problèmes méthodologiques difficiles à résoudre. Dans la plupart des évaluations, on procède par

comparaison d'un groupe expérimental bénéficiant des mesures prévues dans le cadre de ces politiques avec un groupe témoin ne bénéficiant pas de ces mesures. Or, nombreux sont les chercheurs qui reconnaissent que la condition *ceteris paribus* est rarement respectée. Nous-mêmes avons fait face à cette difficulté.

Dans une première étude, nous avons comparé la réussite des élèves scolarisés dans le REP à celle de leurs camarades scolarisés dans l'ensemble des autres établissements genevois (Soussi et al., 2008). Nos résultats ne permettent pas de conclure à des effets positifs du REP. Il demeure un écart non négligeable de réussite en français et en mathématiques des élèves scolarisés dans le REP par rapport à ceux scolarisés dans les autres établissements primaires du canton. Cette absence d'évolution de cet écart peut être attribuée au fait que l'on compare des élèves vivant dans des conditions de vie difficiles avec un échantillon de sujets tout venant, mélangeant enfants de familles favorisées et défavorisées. Sans conteste, ce type de comparaison ne respecte pas le principe de « toutes choses égales par ailleurs ».

Dans une deuxième étude, nous avons comparé les performances des élèves du REP à celles de camarades scolarisés dans des établissements « témoins » dont les caractéristiques sociodémographiques sont similaires (Soussi, Nidegger, Dutrévis & Crahay, 2012). Les analyses multiniveaux (à deux ou à trois niveaux : élèves, établissement ou élèves, classe, établissement) ont surtout mis en évidence le poids des caractéristiques individuelles, mais peu d'effets des établissements et de l'appartenance au REP.

Dans cette deuxième étude-ci, nous nous rapprochons de la comparaison expérimentale. Cependant, les deux groupes ne sont pas parfaitement identiques. Le groupe témoin se compose de six établissements ayant un public d'élèves présentant des caractéristiques sociodémographiques légèrement moins défavorisées par rapport au groupe expérimental. Ce sont d'ailleurs des établissements entrés par la suite dans le REP lorsque les critères ont été élargis. Quant au groupe expérimental, composé (au moment de l'étude) de 15 établissements du REP, il n'est pas homogène, et ce, à différents égards. Ainsi, la condition principale pour qu'un établissement intègre le REP est qu'il accueille une proportion d'enfants de milieux défavorisés d'au moins 55%. Or, ce pourcentage varie de 52% à plus de 70% selon les établissements. Il en va de même pour la proportion d'élèves allophones, critère utilisé de façon complémentaire : elle varie de

40% (moyenne des établissements primaires genevois) à plus de 70%. Il importe également de souligner que les établissements sont entrés dans le REP à des moments différents, ce qui a pour conséquence une durée variable d'exposition des enseignants et des élèves au dispositif.

En effet, le Réseau d'enseignement prioritaire, mis en place dès 2006 à l'école primaire pour favoriser l'égalité des chances en matière de réussite scolaire dans tout le canton, s'est constitué en plusieurs phases :

- En 2006, l'expérience REP a débuté avec une seule école pilote. Lors de cette première phase, les deux principales mesures du REP (réduction du nombre d'élèves par classe et embauche d'un éducateur scolaire) ont été introduites en même temps que différentes autres mesures expérimentées dans le REP, telles que la création de directions, puis de projets et de conseils d'établissement.
- En 2007-2008, huit autres écoles sont entrées dans le REP. Peu à peu, les autorités scolaires ont également décidé de regrouper les 220 écoles en 91 établissements en lien avec la création de directions d'établissement.
- En 2008-2009, huit autres écoles ont intégré le REP, dont certaines qui avaient été regroupées avec les écoles de la phase 2.
- En 2009-2010, un établissement a rejoint le REP.
- En 2010-2011, deux établissements de plus sont entrés dans le REP.
- En 2011-2012, enfin, deux écoles, dont l'une était regroupée avec une autre déjà dans le REP, ont rejoint le réseau, tandis qu'un établissement dont la composition socioéconomique s'était améliorée en est sorti. Le REP est, depuis la rentrée 2011, composé de 17 établissements sur 81. Depuis cette année, les établissements ont fait l'objet de regroupements et ceux du REP représentent 13 établissements sur 58.

L'évaluation des effets « vrais » de l'entrée dans le REP pour un établissement est donc problématique et il n'est pas assuré que la méthode comparative soit la meilleure façon de procéder. Dans cet article, nous explorons une autre voie inspirée de l'analyse des séries temporelles¹ (voir Campbell & Stanley, 1963 ; Shadish, Cook & Campbell, 1971 ; Barlow, Nock & Hersen, 2008).

Méthode

Les analyses dont il est question dans cet article portent sur les établissements² qui composent le REP. Cependant, les établissements se sont constitués au fur et à mesure, et certains sont composés de plusieurs écoles qui ne sont pas entrées dans le dispositif en même temps. C'est pourquoi nous avons considéré qu'il y avait 20 établissements, et non 17, qui composent le REP.

Des analyses multiniveaux de croissance permettant d'évaluer une évolution et d'observer les effets du contexte (notamment ceux des établissements) auraient pu être réalisées puisque les élèves sont soumis à des évaluations externes en français et en mathématiques à trois reprises lors de leur scolarité primaire. Toutefois, nous nous heurtons à la même difficulté que celle décrite précédemment, compte tenu de la constitution progressive du REP de 2006 à 2011. Le REP ne constitue pas un ensemble homogène, car, selon les établissements, les élèves n'ont pas été exposés le même nombre d'années au dispositif.

La première partie de cette étude a pour but d'étudier globalement et par établissement l'évolution des résultats des élèves sur une période de quatre ans (voir les quatre moments dans le Tableau 1).

Tableau 1
Plan expérimental (séries temporelles)

Moment 1	Moment 2	Moment 3	Moment 4
Une année avant l'introduction du REP	Année d'introduction du REP	Une année après l'introduction du REP	Deux ans après l'introduction du REP
t-1	t	t+1	t+2

On retrouve donc un schéma classique du type « avant – traitement – après (2) », partant de l'hypothèse que, si le dispositif REP est efficace, on devrait observer des résultats qui s'améliorent entre t-1 et t+1, et qui se maintiennent (ou continuent de s'améliorer) en t+2.

Dans cette perspective, notre étude s'intéresse tout particulièrement aux résultats des différents établissements, dont on peut notamment étudier les « profils d'évolution » définis par les quatre moyennes observées entre t-1 à t+2. Le profil de chaque établissement est donc exprimé par une série de quatre moyennes ordonnées.

Vu sous cet angle, le problème est tout à fait classique et peut être traité en appliquant la méthode usuelle d'analyse de la variance sur un schéma défini par deux facteurs : le facteur *Moments* (quatre modalités ou niveaux : de t-1 à t+2) et le facteur *Établissements* (20 modalités). De plus, puisque les quatre moments sont ordonnés dans le temps, on peut appliquer au facteur *Moments* des contrastes polynomiaux. Cette approche permettra en particulier d'étudier la tendance (c.-à-d. l'évolution) des résultats observés entre t-1 et t+2 : tendances linéaire, quadratique et, éventuellement, cubique³ (Howell, 1998).

Par ailleurs, outre l'effet des deux facteurs sur les résultats des élèves, il est également possible d'étudier l'effet dû à l'interaction entre ces mêmes facteurs, ce qui permet notamment de vérifier si l'effet du moment dépend de l'établissement (ou, selon le point de vue adopté, si l'effet de l'établissement dépend du moment). Dans le cadre de notre étude, la réponse à cette question est cruciale puisqu'elle montrera si les profils d'évolution des établissements diffèrent plus ou moins les uns des autres. À cet égard, on dira que deux profils sont identiques si les tracés qui les représentent sont parallèles, ou se superposent dans le cas limite (même type d'évolution ; voir les figures de l'interaction avec, en abscisse, les quatre moments). Cette définition s'applique également aux profils tels que nous les considérerons dans la deuxième partie de cet article.

Signalons enfin qu'à partir d'une analyse de la variance globale, il est possible d'envisager l'étude de certains phénomènes plus spécifiques (ou sectoriels) en considérant l'approche dite des comparaisons partielles (ou multiples), par exemple la méthode de Bonferroni, issue de la famille des comparaisons dites *a posteriori* ou post-hoc (Howell, 1998).

Il faut retenir également que ces différents traitements (tout comme ceux que nous présenterons ultérieurement) ont été effectués en considérant les résultats fournis par des épreuves (ou évaluations) externes :

- dans deux disciplines : compréhension de l'écrit (CE) et mathématiques (maths) ;
- et

- dans deux années d'études (en début et en fin de primaire): deuxième année (2P) et sixième année (6P) primaires⁴. Les résultats utilisés ont été exprimés en scores centrés réduits (ou scores z)⁵ par élève, puis nous avons calculé la moyenne de ces scores individuels par établissement. L'utilisation de scores z permet de comparer les résultats à des évaluations différentes.

Notre base de données comportait donc quatre situations différentes: CE en 2P, maths en 2P, CE en 6P et maths en 6P.

Les objectifs de cette première analyse sont donc les suivants: 1) vérifier quelle est l'évolution globale (ensemble des établissements) à travers les moments considérés et 2) vérifier s'il existe une interaction entre les deux facteurs considérés, c'est-à-dire si le résultat des établissements varie en fonction du moment considéré.

Nous avons choisi d'utiliser les résultats aux évaluations externes étant donné que nous avons dû travailler sur des données existantes présentant le mérite d'être communes à tous les élèves d'une année d'études et pour une même année scolaire. De plus, les résultats de ces évaluations entrent dans les notes des élèves et participent à la certification en fin de cycle. Quelques biais potentiels peuvent toutefois être relevés. Ainsi, il faut préciser que ces épreuves sont plus ou moins standardisées en matière de passation et de correction, malgré des consignes très précises. Ajoutons qu'il ne s'agit pas des mêmes évaluations d'une année à l'autre, même si elles sont construites de façon identique, sont basées sur un tableau de spécifications portant sur le plan d'études et mesurent les mêmes objectifs d'une année à l'autre. Il en va de même pour les échantillons d'élèves considérés. Ils proviennent des mêmes établissements, mais ne sont pas strictement équivalents d'une année à l'autre (voir Tableau 2).

Dans un second temps, nous avons essayé d'établir une typologie des évolutions observées et d'en dégager les caractéristiques qui les définissent au moyen d'une analyse de grappes (*cluster analysis*).

Comme nous avons déjà eu l'occasion de le préciser, un des objectifs de cette étude est de vérifier comment évoluent les résultats des élèves, globalement et par établissement, entre les quatre moments considérés. À cet effet, il convient tout particulièrement de déterminer si l'évolution constatée est à peu près la même pour l'ensemble (la majorité) des établissements ou si, au contraire, cette évolution varie beaucoup d'un établissement à l'autre. Poursuivant l'étude dans cette direction, nous avons cru nécessaire

Tableau 2
*Nombre d'élèves dans les établissements en fonction
 de l'année scolaire et du moment*

Établissements	Moments							
	t-1		t		t+1		t+2	
	2P	6P	2P	6P	2P	6P	2P	6P
N*	N	N	N	N	N	N	N	N
1	34	38	31	41	50	33	33	25
2	84	69	67	71	76	74	76	58
3	57	40	47	41	32	40	40	33
4	60	31	66	36	57	34	47	54
5	33	67	32	40	29	29	28	30
6	23	19	21	19	18	23	28	17
7	48	43	47	25	45	38	38	37
8	52	54	38	59	71	49	47	51
9	71	89	62	87	83	77	68	81
10	51	données non disponibles	46	46	40	40	42	45
11	20	24	20	27	15	25	13	13
12	55	53	59	58	40	42	39	46
13	34	39	43	34	35	42	21	31
14	31	34	44	33	38	25	22	29
15	48	40	46	41	64	40	55	34
16	45	44	40	38	36	38	38	40
17	43	47	31	46	42	34	61	58
18	22	27	17	31	23	27	14	20
19	21	42	29	31	20	40	24	34
20	38	40	37	49	39	48	38	45

Note. * = Nombre d'élèves soumis à l'évaluation concernée par établissement. Ce nombre est variable d'un établissement à l'autre dépendant du nombre de classes de l'année d'études concernée. Toutefois, il s'agit non d'un échantillon, mais de la population totale par année d'études de l'établissement. Dans certains cas, les mesures reposent sur un petit nombre de classes et les résultats sont dépendants de la volée concernée.

de chercher à regrouper les établissements qui présentent des profils d'évolution semblables et à distinguer des groupes d'établissements (grappes) caractérisés par des profils bien contrastés. (Nous recherchons des groupes aussi homogènes et aussi différenciés que possible.) À partir de ces groupes, la question est alors de vérifier s'il existe des caractéristiques propres aux établissements et aux élèves qui les fréquentent susceptibles d'«expliquer» les ressemblances et les différences observées.

L'objectif de la démarche est donc d'effectuer une classification – soit d'établir une typologie – des établissements en fonction des profils qui décrivent leur évolution.

De ce point de vue, l'analyse typologique classique (ou classification) est la démarche statistique appropriée, qui sera appliquée à un tableau à double entrée présentant quatre colonnes (une pour chaque moment) et autant de lignes qu'il y a d'établissements. L'analyse typologique peut être définie de la façon suivante :

«Étant donné un ensemble d'objets (ou d'individus) décrits par un certain nombre de caractéristiques (ou variables), constituer des groupes (ou types) d'objets tels que les objets soient les plus similaires possibles au sein d'un groupe et que les groupes soient aussi dissemblables que possible ; la ressemblance ou la dissemblance étant mesurée sur l'ensemble des variables décrivant les objets» (Evrard, Pras & Roux, 2003).

Sur le plan pratique, ce type d'analyse peut être effectué en appliquant différents procédés. Or, il n'est pas toujours facile d'en privilégier un parmi ceux que les logiciels statistiques proposent. Disons simplement que la méthode classique, basée au départ sur le calcul des distances entre chaque établissement et chacun des autres, s'est avérée peu utile. En raison de certaines caractéristiques des données, elle conduit simplement à isoler quelques établissements très différents des autres. Il en résulte que les premiers groupes comportent un ou quelques établissements seulement, dont la caractéristique principale est de se démarquer très nettement de tous les autres. Poursuivre dans cette direction aurait donc conduit à étudier uniquement quelques établissements très atypiques et à ignorer ou à négliger la plupart d'entre eux. Après divers essais d'analyse et d'interprétation (en variant notamment le nombre de groupes demandés), il est apparu que la méthode la plus satisfaisante est celle dite des nuées dynamiques (ou centres mobiles ou encore k-moyennes, désignées k-means dans la littérature et les logiciels anglophones). C'est donc en utilisant cette variante de la méthode que les résultats présentés ici ont été obtenus.

Selon nous, cette démarche devrait permettre de répondre à deux questions :

- 1) Existe-t-il une évolution positive (c.-à-d. des résultats globalement plus élevés) entre l'année -1 (année précédant l'introduction du nouveau dispositif) et l'année 0 (année au cours de laquelle ce dispositif a été appliqué pour la première fois)?
- 2) Comment les résultats des établissements ont-ils évolué au cours des deux années qui ont suivi l'introduction du nouveau dispositif (+1 et +2)?

Résultats

Nous allons présenter les résultats en plusieurs phases : examiner la tendance globale selon les quatre moments considérés, puis en fonction des établissements et des périodes (p. ex., t-1/t ou t/t+2 ou encore t-1/t+2). Nous essaierons ensuite de dégager des profils d'évolution par discipline ou par année d'études (en début ou en fin de primaire). L'objectif sera d'essayer de les expliquer.

De manière globale, l'amélioration entre l'année précédant l'entrée dans le dispositif et deux ans après est assez peu marquée. Comme l'illustre la figure 1, des différences d'évolution sont observées selon la discipline considérée. En effet, même si les progrès sont relativement minimes, un écart positif de score de t-1 à t+2 est relevé pour les mathématiques au début de la scolarité, mais pas à la fin du primaire. Ce n'est pas le cas pour la compréhension de l'écrit, surtout en fin de primaire.

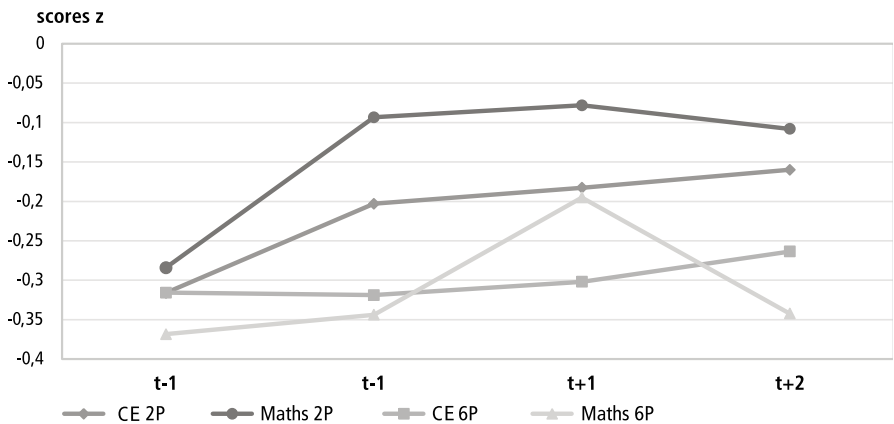


Figure 1. *Évolution des scores z de la période t-1 à la période t+2 en CE et en maths 2P et 6P*

Tableau 3
*Évolution globale des établissements
entre le moment t-1 et le moment t-2*

Type d'évolution	CE		Maths	
	2P	6P	2P	6P
	N ^{bre} d'établissements			
Évolution positive (t-1 < t+2)	11	9	11	4
Stagnation (t-1=t+2)	1	2	2	2
Évolution négative (t-1 > t+2)	8	9	7	14

Le tableau 3 montre une situation très contrastée selon les établissements. Ainsi, pour la compréhension de l'écrit, près de la moitié des établissements (2P = 11/20 et 6P = 9/20) connaissent une évolution positive entre le moment ayant précédé l'entrée dans le dispositif et deux ans après son introduction. En mathématiques, la situation diffère selon l'année d'études : en 2P, 11 établissements sur 20 présentent une évolution positive, tandis que 7 en présentent une négative. À la fin du primaire, la majorité des établissements en maths 6P (14/20) évoluent négativement. Relevons que, dans l'ensemble, de rares établissements ont des résultats identiques entre les deux moments extrêmes (t-1 et t+2), c'est-à-dire que leurs résultats ne semblent pas avoir évolué.

Quand on détaille les évolutions en fonction des établissements, on constate qu'ils se différencient passablement, non seulement en matière d'écart d'un moment à l'autre, mais également de type d'évolution (plutôt ascendante ou plutôt descendante), et ce, aussi bien en compréhension de l'écrit (CE) qu'en mathématiques (maths). Ainsi, en CE 2P (voir Figure 2), l'établissement 6 connaît d'abord une évolution négative puis positive relativement marquée puis à nouveau négative, ce qui correspond à une stagnation entre l'année précédant l'entrée dans le dispositif et deux années après son introduction. L'établissement 4 illustre une évolution positive relativement linéaire. L'établissement 1 évolue également positivement, même s'il suit une progression moins linéaire puisqu'il commence par une évolution négative. Enfin, l'établissement 14 présente relativement peu d'évolution.

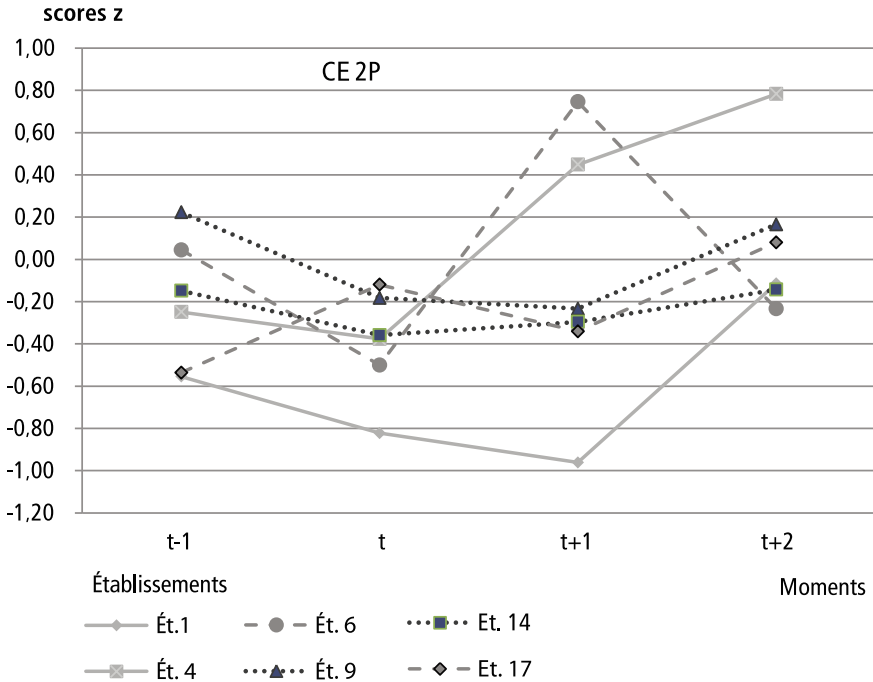


Figure 2. *Évolution par établissement en CE 2P*

À noter que, dans la figure 2, n'apparaît qu'une partie des établissements sélectionnés pour représenter les différents types d'évolution : évolution ascendante en traits pleins, stagnation en pointillés fins ou évolution descendante en pointillés longs.

En maths 6P (voir Figure 3), les profils d'évolution sont également très variables d'un établissement à l'autre. Par exemple, l'établissement 5 connaît une évolution positive relativement régulière. L'établissement 19, qui progresse également, a une évolution moins régulière : il commence par baisser (de t-1 à t), puis progresse. L'établissement 18 a tendance à stagner : il progresse un peu au début, pour retomber, au bout de deux ans, pratiquement au point de départ. Enfin, plusieurs établissements commencent par progresser, puis l'évolution devient négative. C'est le cas de l'établissement 15.

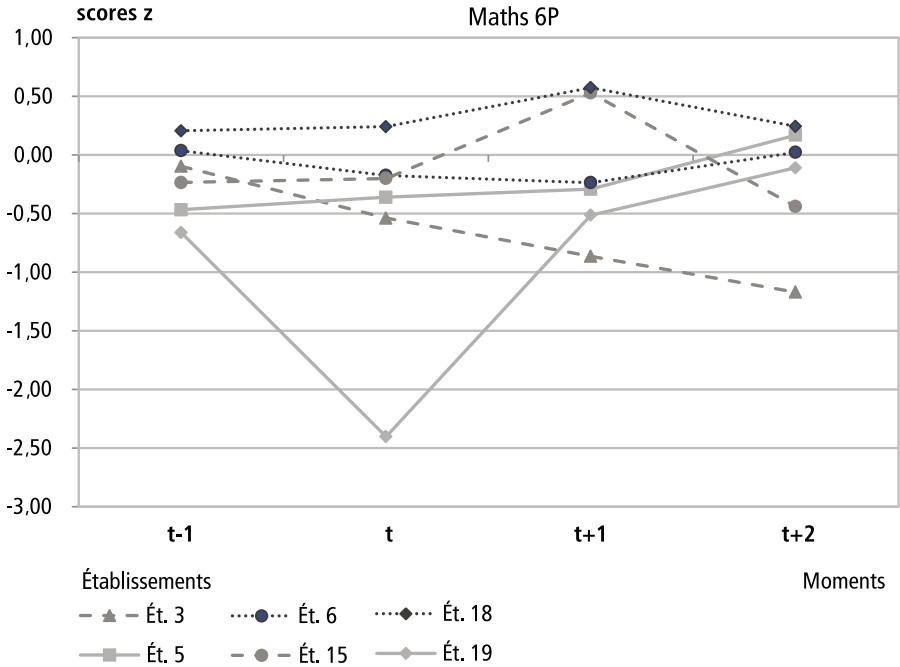


Figure 3. *Évolution par établissement en maths 6P*

Le type d'évolution est d'ailleurs variable selon la discipline et l'année d'études, si l'on observe les contrastes polynomiaux (voir Tableau 4). Ainsi, en compréhension de l'écrit, il n'y a aucune évolution statistiquement significative, pas plus en 2P qu'en 6P. Par contre, en maths 2P, on observe une légère tendance linéaire (qui évolue positivement de manière relativement progressive, passant en moyenne de -0,28 à -0,11) et quadratique (qui évolue positivement entre t-1 et t, puis qui a tendance à stagner entre t et t+1, pour à nouveau augmenter entre t+1 et t+2).

Tableau 4
Contrastes polynomiaux (moments) en CE et en maths 2P et 6P

	CE		Maths	
	2P	6P	2P	6P
<i>Linéaire</i>				
Estimation du contraste	0,109	0,039	0,122	0,051
Significativité	0,164 (n. s.)	0,507 (n. s.)	0,062 ^X	0,629 (n. s.)
<i>Quadratique</i>				
Estimation du contraste	-0,045	0,021	-0,110	-0,086
Significativité	0,563 (n. s.)	0,721 (n. s.)	0,089 ^X	0,413 (n. s.)
<i>Cubique</i>				
Estimation du contraste	-0,021	0,000	0,029	-0,094
Significativité	0,784 (n. s.)	0,996 (n. s.)	0,651 (n. s.)	0,369 (n. s.)

Note. n. s. = n'est pas significatif statistiquement;

X = est tendanciellement significatif (à un seuil de 0,10).

Des évolutions différentes en fonction des deux sous-périodes (t-1 et t; t et t+2)?

Étant donné la grande variété d'évolutions observées pour les quatre moments considérés, nous avons décidé de subdiviser les périodes : d'une part, la période précédant l'entrée dans le dispositif et celle de son entrée, d'autre part, l'année de l'entrée et la période se situant deux ans plus tard. En d'autres termes, y a-t-il des effets de l'introduction du dispositif? Et ces effets sont-ils durables au bout de deux ans?

L'analyse post-hoc réalisée sur les rendements aux différents moments (par le test de comparaison de Bonferroni) met en évidence des évolutions variées selon la discipline et l'année d'études considérées (voir Tableau 5).

Tableau 5
Comparaison des rendements par moment pris 2 à 2

Disciplines	Moments	t-1	t	t+1	t+2	t-1	t	t+1	t+2
CE 2P	t	n. s.	t-1	n. s.	t-1	n. s.	t-1	n. s.	n. s.
	t+1	n. s.	t+1	n. s.	t	n. s.	t	n. s.	n. s.
	t+2	n. s.	t+2	n. s.	t+2	n. s.	t+1	n. s.	n. s.
Maths 2P	t	X	t-1	X	t-1	*	t-1	X	X
	t+1	*	t+1	n. s.	t	n. s.	t	n. s.	n. s.
	t+2	X	t+2	n. s.	t+2	n. s.	t+1	n. s.	n. s.
CE 6P	t	n. s.	t-1	n. s.	t-1	n. s.	t-1	n. s.	n. s.
	t+1	n. s.	t+1	n. s.	t	n. s.	t	n. s.	n. s.
	t+2	n. s.	t+2	n. s.	t+2	n. s.	t+1	n. s.	n. s.
Maths 6P	t	n. s.	t-1	n. s.	t-1	n. s.	t-1	n. s.	n. s.
	t+1	n. s.	t+1	n. s.	t	n. s.	t	n. s.	n. s.
	t+2	n. s.	t+2	n. s.	t+2	n. s.	t+1	n. s.	n. s.

Note. n. s. = n'est pas significatif statistiquement ;
 X = est tendanciellement significatif (à un seuil de 0,10) ;
 * = est significatif à un seuil de 0,05.

Globalement, il ressort qu'il y a très peu d'effets entre le moment précédant l'introduction du REP et l'année d'après, à l'exception des mathématiques au début de la scolarité : le score z moyen va en augmentant, passant de -0,28 à -0,09. Il en va de même pour la période allant de l'introduction du dispositif à deux ans plus tard. De nouveau, en maths 2P seulement, une différence positive et statistiquement significative entre le moment t et t+2 est observée : le score z passe de -0,28 à -0,11.

Malgré ces résultats peu encourageants et étant donné les évolutions très variables d'un établissement à l'autre, nous avons décidé de poursuivre les analyses pour essayer de comprendre ce qui pouvait expliquer une évolution positive ou, au contraire, négative.

Les profils d'évolution

Nous avons d'abord réalisé une analyse de grappes pour chaque discipline et chaque année d'études. Pour chaque analyse, nous indiquerons dans quelle mesure les grappes regroupent ou non des établissements ayant les mêmes caractéristiques sociodémographiques. Nous avons choisi une solution en trois grappes (ou en deux, p. ex. pour les maths 2P) pour deux raisons : la première étant qu'avec un nombre relativement petit d'établissements, elle présente l'avantage que chaque grappe comporte plusieurs établissements, la seconde étant que l'évolution observée à l'intérieur d'une grappe soit comparable.

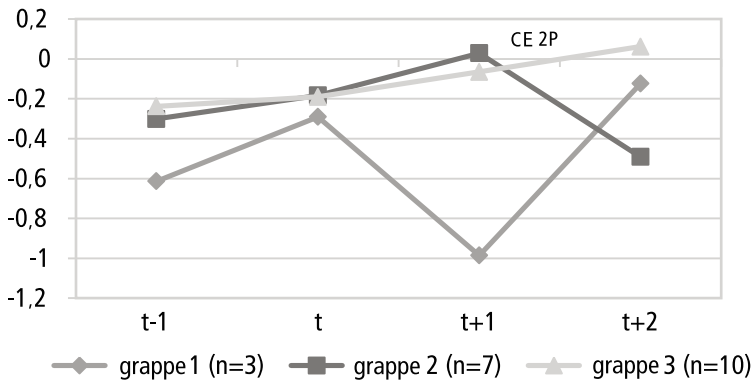


Figure 4. *Profil d'évolution en CE 2P*

En CE 2P, les trois grappes connaissent des profils d'évolution différents (voir Figure 4) :

- La première grappe (3 établissements), dont les résultats sont plutôt bas au départ, présente une évolution positive ;
- La deuxième grappe (7 établissements) présente une évolution globalement négative ;
- La troisième grappe (10 établissements), dont les scores sont les plus élevés, connaît une évolution globalement positive.

Il est difficile d'expliquer par les caractéristiques sociodémographiques des établissements leurs différentes évolutions. Il faut sans doute rechercher d'autres raisons telles que les pratiques pédagogiques, la dynamique ou le climat existant à l'intérieur de l'établissement.

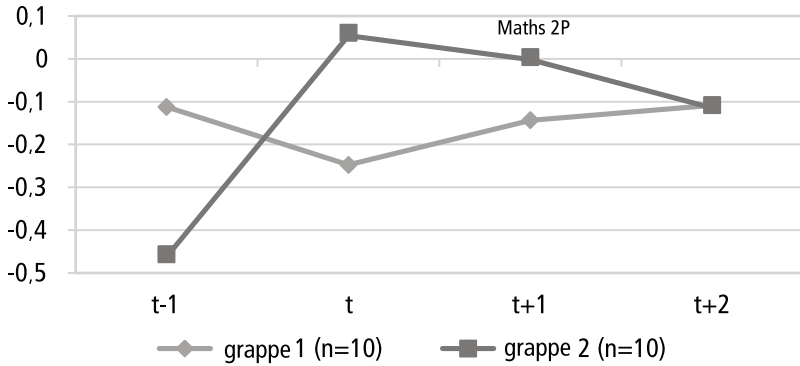


Figure 5. *Profil d'évolution en maths 2P*

En maths 2P, la meilleure solution pour regrouper les établissements est celle en deux grappes comportant chacune la moitié des établissements (voir Figure 5):

- La première grappe (10 établissements) présente des scores proches de la moyenne et a tendance à stagner ;
- La seconde grappe (les 10 autres établissements) commence avec des scores plus faibles, évolue positivement de façon marquée au début, mais retombe un peu lors de la dernière période. L'évolution est globalement positive.

Cette fois, des différences dans le taux d'élèves en retard sont observées, la première grappe comportant une proportion plus élevée d'élèves en retard que la seconde.

À la fin de l'école primaire, la situation se présente un peu différemment du point de vue des établissements qui se regroupent. Pour la compréhension de l'écrit, elle se présente comme suit :

- La première grappe (6 établissements) a tendance à stagner. Ses résultats pour les deux moments extrêmes se situent autour de la moyenne ;
- Dans la deuxième grappe (7 établissements), l'évolution est positive, avec un profil de type cubique. Le score est relativement bas au départ et descend encore au moment t+1. La progression a lieu entre t+1 et t+2 ;
- La troisième grappe (7 établissements) se caractérise par une évolution négative linéaire.

Il est difficile d'expliquer les raisons de ces regroupements, hormis le fait que les établissements faisant partie d'une grappe évoluent de la même manière: pas plus l'entrée dans le REP que la composition sociodémographique ne se différencient particulièrement d'une grappe à l'autre. Il faut sans doute rechercher des raisons liées à la dynamique des établissements ou aux pratiques pédagogiques utilisées.

En mathématiques, plusieurs solutions ont été testées. Celle à trois grappes a été finalement retenue, même si la deuxième grappe ne comporte qu'un établissement. La situation se présente de la manière suivante: une première grappe se compose de 11 établissements, dont les résultats sont relativement bons et ont tendance à stagner. La deuxième grappe ne se compose que d'un seul établissement atypique, avec des résultats très faibles. La troisième grappe regroupant 8 établissements présente une évolution négative. Les différentes grappes ne se différencient pas sur le plan des caractéristiques sociodémographiques.

Compte tenu des résultats mis en évidence dans le tableau 5, à savoir une certaine similarité à l'intérieur d'une discipline, nous avons décidé de réaliser une seconde analyse de grappes par discipline (compréhension de l'écrit d'une part et mathématiques d'autre part) afin d'observer s'il y aurait davantage de logique de regroupement à l'intérieur des établissements.

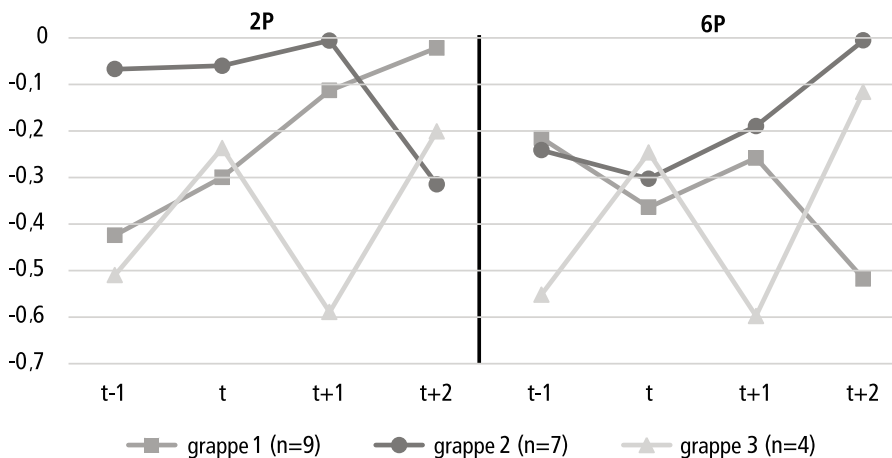


Figure 6. Profil d'évolution en CE 2P et 6P

L'analyse de grappes réalisée par discipline pour les deux années scolaires (début de primaire et fin de primaire) montre les résultats suivants. Pour la compréhension de l'écrit (voir Figure 6), la situation se présente comme suit avec trois grappes :

- Une première grappe (9 établissements) connaît une évolution positive assez marquée en début de primaire, mais une évolution négative en fin de primaire ;
- Dans la deuxième grappe (7 établissements), la situation est inversée : l'évolution est négative en début de primaire et positive en fin de primaire ;
- Enfin, la troisième grappe (4 établissements) se caractérise par une évolution très similaire au début et à la fin du primaire de type cubique. Au départ, les scores augmentent, baissent, puis remontent, pour se stabiliser à un niveau un peu plus élevé qu'au départ.

Quand on observe les caractéristiques des établissements composant ces trois grappes en termes sociodémographiques (proportion d'élèves de milieux défavorisés, d'allophones et d'élèves en retard), on ne constate pas de différences statistiquement significatives, même si la première grappe comporte en moyenne une plus grande proportion d'élèves de milieux défavorisés (59% c. 56% pour la deuxième et 53% pour la troisième). Il faut sans doute chercher ailleurs l'explication de ces similitudes et de regroupement dans les grappes. L'hypothèse suivante peut être avancée : les établissements qui connaissent la même évolution positive en début et en fin de 6P ont particulièrement mis l'accent sur l'apprentissage de la langue tout au long de la scolarité primaire⁶. Dans les autres situations, nous pouvons supposer qu'au contraire il y a une différence de centration entre le début et la fin de la scolarité primaire.

Pour les mathématiques (voir Figure 7), nous avons également adopté une solution en trois grappes. En 2P comme en 6P, la première grappe (2 établissements) connaît des évolutions positives. En 2P, les résultats sont au départ assez bas, puis augmentent. En 6P, ils se situent autour de la moyenne et terminent un peu au-dessus. La deuxième grappe (3 établissements) a un fonctionnement plus différencié entre les deux moments de la scolarité : en 2P, elle a une tendance à la baisse, alors que c'est l'inverse en 6P. Enfin, la troisième grappe, qui regroupe une majorité d'établissements (15), présente globalement une tendance à la baisse.

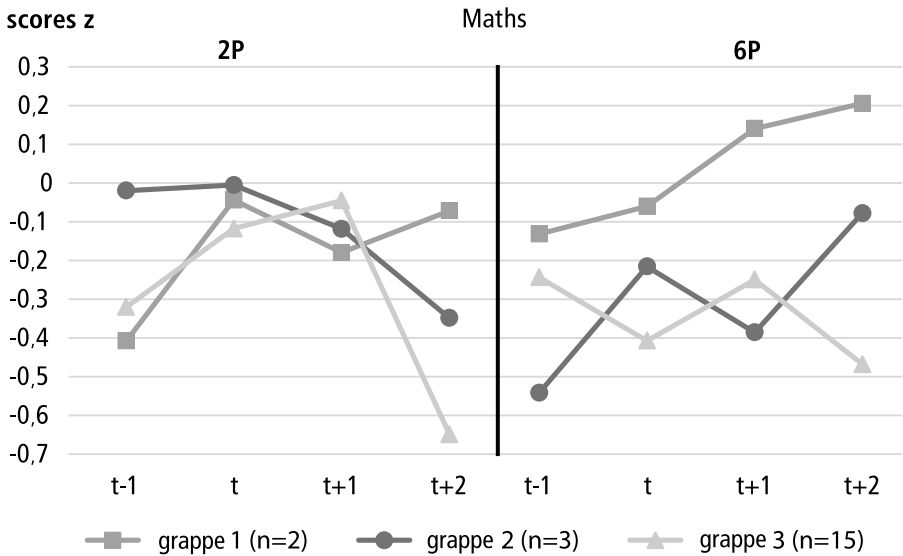


Figure 7. *Profil d'évolution en maths 2P et 6P*

Ces grappes se différencient cette fois de manière importante par la proportion d'élèves provenant de milieux défavorisés : 44% dans la première grappe (c.-à-d. en dessous de la proportion requise pour entrer dans le REP), comparativement à 56% dans la deuxième et 58% dans la troisième.

Compte tenu des évolutions relativement différentes selon la discipline, nous ne présenterons pas les résultats en fonction des années d'études.

Discussion et conclusion

La mesure des effets de dispositifs visant l'égalité des chances et la lutte contre l'échec scolaire est problématique. Le modèle expérimental est rarement applicable tel quel, l'équivalence des groupes étant difficile à obtenir. Rappelons les débats méthodologiques qui, concernant les recherches sur le redoublement, ont pour objet d'assurer l'équivalence « vraie » des doublants et non-doublants (Grisay, 2001 ; Crahay, 2004). Dans le cas du REP genevois, la difficulté à mesurer les effets est d'autant plus grande qu'il s'agit d'un dispositif constitué petit à petit, et comportant un certain nombre de mesures expérimentées dans le dispositif et généralisées

par la suite dans l'ensemble des établissements genevois de l'école primaire. Or, connaissant les effets de testing et, surtout, de connaissance des tests, nous craignons que la diffusion de ces informations n'ait eu un effet sur les pratiques courantes dans les écoles hors du REP.

Les analyses globales comparant les établissements du REP et les autres établissements primaires du canton de Genève n'ont pas pu mettre en évidence d'effets positifs de ce dispositif sur les acquis des élèves observés au moyen des résultats aux épreuves externes. Il faut souligner que ces premières comparaisons portaient sur des établissements trop différents du point de vue des compétences. C'est pourquoi, dans l'étude de Soussi et ses collaborateurs (2012), les auteurs ont tenté, dans un deuxième temps, de comparer les établissements bénéficiant des mesures REP avec d'autres établissements présentant des caractéristiques similaires, mais ne recevant pas de mesures. La réflexion critique sur la recherche oblige à reconnaître qu'ils n'ont pas réussi à respecter la condition *ceteris paribus*, notamment parce que l'histoire de chaque entrée dans le REP est singulière, qu'elle ne se passe pas au même moment et que les motivations pour y entrer peuvent différer.

C'est pourquoi il nous paraissait important de dépasser ce constat et de trouver une manière de mesurer des effets en prenant en compte le temps d'exposition au dispositif, mettant de côté le moment de l'entrée dans le dispositif. La méthode des séries temporelles (en prenant en compte les différents moments considérés) permet de réaliser cet objectif, ce qui n'était pas possible avec d'autres méthodes, par exemple l'analyse multi-niveaux de croissance, pour laquelle il est nécessaire de disposer de plusieurs observations simultanées. Nous avons donc construit un plan expérimental comportant quatre moments définis pour chaque établissement indépendamment de son moment d'entrée dans le REP afin d'observer l'évolution générale des résultats en compréhension de l'écrit et en mathématiques en début et en fin de primaire, depuis l'année précédant l'entrée dans le dispositif et deux ans après son introduction. Nous avons également voulu voir quelle était l'évolution immédiate à court terme et aussi à moyen terme, à savoir au bout de deux ans d'introduction du dispositif, c'est-à-dire si cette évolution se poursuivait.

Le plan expérimental ainsi constitué a permis de montrer des effets d'interaction moments-établissement, c'est-à-dire que chaque établissement présente une évolution spécifique entre les deux moments extrêmes

(celui qui précède l'entrée dans le dispositif et celui qui se situe deux ans après le début de l'expérience). Soulignons que l'évolution est positive dans certains établissements. Ces résultats confirment la variabilité des établissements du REP au-delà de leur variabilité en termes de composition sociodémographique.

Un découpage de l'évolution en deux phases ($t-1/t$ et $t/t+2$) permet de constater que les effets sont également faibles, voire inexistantes dans ces deux phases, à l'exception des mathématiques au début de la scolarité. Ainsi, les résultats observés une année avant l'entrée dans le dispositif se différencient de ceux observés lors de l'entrée dans le dispositif ou encore une année, voire deux années plus tard.

Par ailleurs, les évolutions varient également selon la discipline observée : la compréhension de l'écrit et les mathématiques. Ainsi, une évolution globalement positive est principalement observée en mathématiques en début de primaire pour les phases, mais pas à la fin du primaire.

Cette absence de progrès en compréhension de l'écrit peut étonner en partie étant donné l'accent mis sur la maîtrise de la langue dans la plupart des projets d'établissement. Toutefois, on sait que, dans certains quartiers comportant beaucoup d'élèves de milieux défavorisés et surtout allophones, la maîtrise de la langue est parfois moins grande, ce qui explique des résultats moins bons en lecture. Or, la langue est un vecteur de connaissances scolaires globales et une maîtrise insuffisante a des répercussions dans toutes les disciplines. C'est pourquoi des efforts sont souvent mis pour développer les compétences langagières.

Par ailleurs, une évolution positive globale est observée en mathématiques au début du primaire, mais pas en fin de primaire. Cela pourrait sans doute s'expliquer par deux éléments. Le premier élément concerne le caractère cumulatif des connaissances en mathématiques. Il est certainement difficile de rattraper l'écart de départ. De plus, on peut supposer que plus les élèves avancent dans la scolarité, plus ils font face à des problèmes complexes qui les discriminent davantage et qui peuvent avoir un effet négatif sur les élèves du REP (voir Soussi et al., 2008 ; Bautier & Goigoux, 2004 ; Cèbe & Goigoux, 2004), notamment avec la gestion des implicites et les processus de secondarisation nécessaires à la construction des apprentissages scolaires. De manière générale, l'origine socioéconomique

influe sur les compétences en mathématiques aussi bien que sur celles en lecture (voir les études du programme PISA, notamment OCDE, 2014; Andreu, Ben Ali & Rocher, 2016).

Le second élément est sans doute lié au fait du temps d'exposition au REP de la part des élèves. En effet, les plus jeunes élèves fréquentant des établissements du REP ont commencé leur scolarité dans le dispositif, ce qui n'est pas le cas de ceux en fin de primaire, qui ont été exposés moins longtemps à ce dispositif.

Pour essayer de dégager des tendances et d'expliquer ce qui détermine le type d'évolution, nous avons réalisé une analyse de grappes. Il est toutefois difficile de trouver une logique de regroupement des établissements sur la base des données récoltées et des caractéristiques sociodémographiques des établissements (proportion d'élèves de milieux défavorisés, d'allophones ou d'élèves en retard) ou encore de la taille de l'établissement ou du moment d'entrée dans le REP. Dans quels cas observe-t-on des progressions ou au contraire des stagnations, voire des baisses depuis le début du dispositif? D'autres raisons sont sans doute à chercher ailleurs, par exemple dans la dynamique à l'intérieur des établissements et des équipes enseignantes ou dans les accents spécifiques mis dans le projet d'établissement (notamment, la lecture) qui pourraient expliquer des évolutions positives dans ce domaine. Des expériences de politiques d'éducation prioritaire menées dans d'autres pays, par exemple en France, ont montré que certaines zones d'éducation prioritaire (ZEP) réussissaient mieux que d'autres (voir Moisan & Simon, 1997). Il faut souligner que le plan expérimental utilisé (séries temporelles) a permis de montrer certaines évolutions positives, ce qui n'avait pas pu être montré avec des analyses quantitatives plus complexes, notamment multiniveaux. Ces séries temporelles offrent une mesure pertinente de la variable dépendante, car elles permettent de prendre en compte le fait que le REP s'est constitué en plusieurs phases et qu'il ne constitue pas un ensemble homogène, les élèves y étant scolarisés n'ayant pas été exposés de façon homogène au dispositif selon l'établissement. Toutefois, on ne dispose pas de mesures précises du changement dont on espère des effets sur les compétences des élèves.

Par ailleurs, nous constatons que les différentes grappes ne regroupent pas forcément les mêmes établissements selon l'année d'études ou la discipline observées. L'évolution est donc à la fois spécifique des établissements, de la discipline considérée, voire de l'année d'études.

Le plan expérimental utilisé dans cette étude a toutefois permis de dépasser un constat peu rassurant qui concluait à une absence d'effets du dispositif REP et de mettre en évidence l'existence de progrès dans certains établissements et en mathématiques, en particulier. Nous croyons qu'il faudrait creuser plus ou autrement le problème et notamment s'intéresser au type de mesures mises en place.

Ainsi, sans prendre en compte l'ensemble des neuf qualités transversales selon le modèle de Bouchard et Plante (2002, cités par Demeuse et Friant, 2012), qui permet d'évaluer la qualité d'une action, nous en retiendrons quelques-unes, parmi lesquelles la pertinence (les objectifs répondent-ils aux besoins?), l'efficacité et l'efficience. Ainsi, le REP, qui a pour objectif de lutter contre l'échec scolaire et de promouvoir l'égalité des chances, comporte assez peu de mesures : dans des établissements socialement défavorisés, le nombre d'élèves par classe est diminué de deux et on a créé la fonction d'éducateurs scolaires permettant de faire le lien entre la famille et l'école ainsi que l'environnement extérieur. Si les mesures mises en place visent à améliorer les conditions d'enseignement, peu de mesures pédagogiques ou didactiques ont été prévues sur le plan institutionnel. Cela ne signifie pas que les équipes enseignantes ne s'adaptent pas à leur public. Mais ces résultats, comme ceux obtenus dans les études précédentes, tendent à montrer la nécessité d'aller plus loin dans l'évaluation et l'observation des effets. Il serait nécessaire d'observer de façon plus fine et plus qualitative la mise en œuvre concrète du REP, en particulier dans les établissements qui connaissent une évolution positive, pour essayer d'en comprendre les raisons, notamment du point de vue des pratiques pédagogiques et de la dynamique de l'établissement.

Réception : 31 août 2015

Version finale : 23 janvier 2017

Acceptation : 7 février 2017

NOTES

1. Une série temporelle (ou encore une série chronologique) est une suite finie (x_1, \dots, x_n) de données indexées par le temps. L'indice temps peut être, selon les cas, la minute, l'heure, le jour, l'année, etc. Le nombre n est appelé la longueur de la série.
2. Dans le cadre de la réorganisation de l'école primaire, des directions, des projets et des conseils d'établissement ont été instaurés. Les écoles se sont également regroupées en établissements. Dans certains cas, les établissements composés de plusieurs écoles ne sont entrés que partiellement dans le REP (p. ex., une école sur deux), étant donné que l'une des écoles ne remplissait pas les conditions.
3. Une droite de tendance linéaire (de premier degré) est exprimée par une équation de type $Y = aX + b$. L'évolution est de type quadratique (de second degré: $Y = aX^2 + \dots + b$) lorsqu'elle connaît un changement de direction (p. ex., dans un premier temps, elle est ascendante et, dans un second temps, elle est descendante ou l'inverse). Elle est de type cubique (ou de troisième degré: $Y = aX^3 + \dots + b$) quand elle connaît deux changements de direction.
4. En deuxième année primaire (actuellement 4P HarmoS depuis l'entrée en vigueur du processus d'harmonisation scolaire en Suisse caractérisé notamment par l'entrée à l'école obligatoire à 4 ans), les élèves ont entre 7 et 8 ans. En fin de primaire (actuellement 8P HarmoS), les élèves ont entre 11 et 12 ans.
5. Le score z (ou score centré réduit) exprime l'écart par rapport à la valeur moyenne, en écart-type. Sa moyenne est de 0 et son écart-type de 1. Un score z de 0 indique qu'on se trouve à la moyenne, tandis qu'un score de 1 indique qu'on est à un écart-type au-dessus de la moyenne. S'il atteint 2 ou 3, on se trouve dans une zone très supérieure, alors que, s'il se trouve à -2 ou -3, on est dans la zone très inférieure. L'utilisation de scores z permet de comparer des résultats dont les notes brutes n'ont pas les mêmes moyennes ou écarts-types.
6. L'analyse des projets d'établissement a montré des différences dans les objectifs ciblés (voir Gros, Guillely, Jaeggi & Sermey, 2012; Soussi & Nidegger, 2010).

RÉFÉRENCES

- Andreu, S., Ben Ali, L. et Rocher, T. (2016). Évaluation numérique des compétences du socle en début de sixième : des niveaux de performance contrastés selon les académies, *Note d'information no 18*. Paris : Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Direction de l'évaluation, de la prospective et de la performance.
- Barlow, D. H., Nock, M. K., & Hersen, M. (2008). *Single case research designs: Strategies for studying behavior change*. New York, NY: Allyn & Bacon.
- Bautier, E. et Goigoux, R. (2004). Difficultés d'apprentissage, processus de secondarisation et pratiques enseignantes : une hypothèse relationnelle. *Revue française de pédagogie*, 142, 89-100. Repéré à www.persee.fr/doc/rfp_0556-7807_2004_num_148_1_3252
- Bouchard, C. et Plante, J. (2002). La qualité : mieux la définir pour mieux la mesurer. *Cahiers du Service de pédagogie expérimentale*, 11-12, 219-236. Repéré à www.fastef-portedu.ucad.sn/cesea/comfr/ulg/cahiers11_11.pdf
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.
- Cèbe, S. et Goigoux, R. (2004). Quelles pratiques d'enseignement pour compenser les inégalités scolaires en grande section de maternelle? Dans L. Talbot (dir.), *Pratiques d'enseignement et élèves en difficulté* (pp.112-121). Toulouse : Érès.
- Crahay, M. (2004). Peut-on conclure à propos des effets du redoublement? *Revue française de pédagogie*, 148(1), 11-23. doi: 10.3406/rfp.2004.3246
- Demeuse, M., Frandji, D., Greger, D. et Rochex, J.-Y. (dir.). (2008). *Les politiques d'éducation prioritaire en Europe: conceptions, mises en œuvre, débats* (tome I). Lyon : Institut national de recherche pédagogique.
- Demeuse, M., Frandji, D., Greger, D. et Rochex, J.-Y. (2011). *Les politiques d'éducation prioritaire en Europe: quel devenir pour l'égalité scolaire?* (tome II). Lyon : ENS Éditions
- Demeuse, M. et Friant, N. (2012). Évaluer les politiques d'éducation prioritaire en Europe : un défi méthodologique. *Revue suisse des sciences de l'éducation*, 34, 39-55. Repéré à <https://hal.archives-ouvertes.fr/hal-00722916/document>
- Dutrévis, M. et Crahay, M. (2009). Les politiques de discrimination positive : un bilan impossible. Dans G. Chappelle et M. Crahay (dir.), *Réussir à apprendre* (pp. 59-73). Paris : Presses universitaires de France.
- Evrard, Y., Pras, B. et Roux, E. (2003). *Market: études et recherche en marketing*. Paris : Nathan.
- Grisay, A. (2001). Évaluer les dispositifs de prise en charge d'élèves faibles (ou forts) : l'utilisation de groupes naturels entraîne des artefacts. *Cahiers du Service de pédagogie expérimentale*, 7-8, 129-151. Repéré à www.fastef-portedu.ucad.sn/cesea/comfr/ulg/cahiers7_10.pdf
- Gros, D., Guilley, E., Jaeggi, J.-M. et Sermet, G. (2012). *Évaluation externe des projets d'établissement de l'enseignement primaire*. Rapport de synthèse. Genève: DIP, SRED.
- Howell, D. C. (1998). *Méthodes statistiques en sciences humaines*. Bruxelles: De Boeck.

- Moisan, C. et Simon, J. (1997). *Les déterminants de la réussite scolaire en zone d'éducation prioritaire*. Paris : Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Inspection générale de l'éducation nationale.
- Organisation de coopération et de développement économiques (OCDE). (2014). *Résultats du PISA 2012 : l'équité au service de l'excellence – Offrir à chaque élève la possibilité de réussir* (vol. II), coll. PISA. Paris : Éd. OCDE. doi: 10.1787/9789264205321-fr
- Rochex, J.-Y. (2008). Les politiques d'éducation prioritaire en Europe, d'un « âge » et d'un pays à l'autre. Dans M. Demeuse, D. Frandji, D. Greger et J.-Y. Rochex (dir.), *Les politiques d'éducation prioritaire en Europe : conceptions, mises en œuvre, débats* (pp. 409-451). Lyon : Institut national de recherche pédagogique.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (1971/2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA : Houghton Mifflin.
- Soussi, A., Guignard, N., Hayoz, E. et Nidegger, C. (2008). *Mise en place du REP : effets sur les compétences des élèves*. Genève : DIP, SRED.
- Soussi, A. et Nidegger, C. (2010). *Le réseau d'enseignement prioritaire à Genève : quels effets sur les acquis des élèves après quelques années ?* Genève : DIP, SRED.
- Soussi, A. et Nidegger, C. (2015). *Le réseau d'enseignement prioritaire genevois (REP) 8 ans après son introduction*. Genève : DIP, SRED.
- Soussi, A., Nidegger, C., Dutrévis, M. et Crahay, M. (2012). Un réseau d'enseignement prioritaire dans le canton de Genève : quels effets sur les élèves ? *Revue française de pédagogie*, 178, p. 53-66. doi: 10.4000/rfp.3548