

Impact des établissements sur les performances des élèves. Une étude basée sur les résultats de l'enquête COFO 2016

Franck Petrucci, Service de la recherche en éducation du canton de Genève (SRED) Genève et Centro competenze innovazione e ricerca sui sistemi educativi (CIRSE) Locarno

Alice Ambrosetti et Francesca Crotta, CIRSE Locarno

Christian Nidegger, SRED Genève

Selon le courant du school effectiveness, les résultats des élèves dépendent de leurs caractéristiques individuelles et du contexte de scolarisation. Dans quelle mesure l'enquête COFO, évaluation standardisée à large échelle réalisée au niveau suisse, peut-elle apporter aux cantons des informations pour juger de l'efficacité de leurs établissements ? Au Tessin et à Genève, la part des différences de résultats liées à l'établissement apparaît extrêmement modeste. Le calcul de la valeur ajoutée montre que la grande majorité des établissements obtiennent des résultats conformes à ceux auxquels on peut s'attendre. Seuls quelques-uns ont des résultats plus faibles qu'escompté mais aucun lien entre leur sous-performance et la composition de leur population scolaire ne peut être établi.

Introduction

Depuis quelques dizaines d'années, de nombreux systèmes éducatifs se sont dotés d'outils de mesure des compétences des élèves. Ainsi, ils peuvent, dans une perspective de vérification et de régulation, s'assurer que les objectifs de connaissances et de compétences qu'ils ont fixés sont atteints. Le processus d'évaluation des performances de l'école sur la base de mesures de performances des élèves est de plus en plus répandu et l'obligation de rendre des comptes (*school accountability*) est devenue un élément central des politiques éducatives dans bon nombre de pays (Loeb et Figlio, 2011). Avec l'arrivée des enquêtes PISA (Programme for International Student Assessment) cette tendance s'est renforcée. En Suisse, suite à PISA 2000, la CDIP (Conférence suisse des directeurs cantonaux de l'instruction publique) a émis un certain nombre de recommandations visant à améliorer les performances des systèmes scolaires cantonaux (CDIP, 2003). L'une d'elles proposait de définir des standards nationaux de compétences devant être atteints par tous les élèves. Ils ont été adoptés en 2011 et font désormais partie intégrante du concordat HarmoS qui vise l'harmonisation de la scolarité obligatoire en Suisse (CDIP, 2007).

La première mesure des standards nationaux a eu lieu en 2016 auprès des élèves de 11^e année pour les mathématiques. Les premières analyses se sont focalisées sur la proportion d'élèves qui atteignent ces compétences fondamentales à l'échelle nationale, point central de l'enquête (Consortium COFO, 2019). Elles ont mis en évidence des différences sensibles entre les cantons et ont montré que les caractéristiques individuelles des élèves et le type de programme suivi n'expliquent que partiellement ces variations. Ces analyses ont, en revanche, peu abordé la question de la variation des résultats à l'intérieur des cantons alors même que PISA a montré que c'est là que se situe la plus grande partie des différences entre élèves les plus faibles et les plus performants (Consortium PISA.ch, 2011). Par ailleurs, il existe désormais un relatif consensus sur l'importance des facteurs contextuels dans l'explication des variations d'acquisitions de compétences. C'est notamment pour cette raison que certaines évaluations standardisées à large échelle collectent en plus des estimations sur les performances une multitude de renseignements sur les élèves, leurs familles, leurs enseignants et leurs écoles, autant d'éléments qui doivent conduire à une meilleure compréhension de l'impact des politiques et des pratiques sur les résultats des élèves (Rutkowski et al., 2014). En tant qu'instrument de monitoring des systèmes scolaires cantonaux, l'enquête COFO (Vérification de l'atteinte des Compétences Fondamentales) doit aider à mieux comprendre ce qui se joue à l'intérieur de chacun d'eux, tout particulièrement du point de vue du lien entre contexte de scolarisation et acquisitions des compétences. Dans ce cadre, l'analyse des effets établissements revêt une importance particulière puisque, selon Grisay (2006), elle constitue l'un des indicateurs les plus distinctifs pour décrire la physionomie d'un système éducatif.

Cette étude constitue une première tentative de mesure des effets établissements à partir des données COFO. Cet instrument national permet-il de se focaliser sur ces entités dont la pertinence en matière d'analyse des variations d'acquisition de compétences a déjà été démontrée (Dumay et Dupriez, 2009) ? Peut-il apporter aux cantons des informations pour juger de l'action de leurs établissements et, en particulier, de leur degré d'efficacité c'est-à-dire de leur capacité à élever le niveau de leurs élèves ? Existe-t-il des établissements plus efficaces que d'autres faisant parvenir leurs élèves à un niveau d'acquisitions supérieur à celui auquel parviennent des élèves comparables mais scolarisés dans d'autres structures (Bressoux, 2006) ? Nous tenterons de répondre à ce type de questions en traitant le cas particulier des cantons du Tessin et de Genève retenus ici pour trois raisons. Tout d'abord, tous leurs établissements de l'enseignement secondaire I accueillent tous les élèves indépendamment de leurs résultats scolaires (seul le lieu de résidence détermine leur affectation). Ceci autorise le traitement de la problématique de l'effet établissement en dépit de la taille limitée des échantillons cantonaux. Là où les élèves sont orientés en fonction de leurs différences d'aptitudes vers des établissements distincts offrant des programmes différenciés il faut, en revanche, procéder à une analyse par filière ce qui est

impossible en raison de la faiblesse des effectifs à disposition. D'autre part, au Tessin et à Genève tous les établissements ont participé à l'enquête ce qui, dans une perspective de monitoring, permet potentiellement d'étudier chaque entité, contrairement à ce qui peut être envisagé dans des cantons plus grands où les établissements ont été échantillonnés pour la réalisation de l'enquête. Enfin, chacun de ces cantons possède une base de données scolaire qui complète les données COFO, notamment en termes de résultats antérieurs des élèves et de caractéristiques des établissements. En effet, contrairement à PISA qui dispose également d'un questionnaire adressé aux chefs d'établissements pour recueillir des renseignements sur l'établissement fréquenté par chaque élève, l'enquête COFO ne collecte que des informations relatives aux élèves (Consortium COFO, 2019), ce qui limite grandement les possibilités d'étude d'un éventuel effet établissement. Compléter les données initialement collectées apparaît donc nécessaire.

Cadre théorique: L'efficacité dans l'enseignement et les établissements scolaires

Depuis les années 60, l'intérêt suscité par l'effet du contexte scolaire sur les performances des élèves n'a cessé de s'accroître (Bressoux, 1994). De très nombreuses recherches ont essayé de voir dans quelle mesure l'apprentissage d'un élève est influencé par le contexte scolaire dans lequel ce dernier est immergé, qu'il s'agisse de l'établissement ou de la classe (Dumay et Dupriez, 2009; Thrupp, 1999). D'une manière générale, ce courant de recherche dit du *school effectiveness* vise à identifier des contextes plus formateurs que d'autres et les facteurs qui y sont associés. Comme l'a indiqué Duru-Bellat (2001) il s'agit d'évaluer la notion de valeur ajoutée, c'est-à-dire «ce que des élèves initialement comparables (mêmes caractéristiques sociales, même niveau scolaire...) «gagnent» ou «perdent» spécifiquement quand ils sont scolarisés dans tel ou tel contexte» (p. 322). L'idée principale est ici de mettre en évidence des effets «nets» du contexte estimés en comparant des élèves au background similaire.

Le courant du *school effectiveness* s'est développé dans les années 1970 en réponse aux travaux de Coleman et al. (1966) qui ont été les premiers à s'intéresser au contexte d'apprentissage, et plus particulièrement à l'établissement. Se situant dans une approche de type *input-output* et traitant largement les écoles comme des «boîtes noires» ils étaient parvenus à la conclusion que l'établissement fréquenté ne faisait pas de différence, attestant d'une certaine incapacité de l'école à rendre la réussite scolaire moins dépendante de l'origine sociale des élèves (Bressoux, 2006). En opposition à ces travaux, les recherches du *school effectiveness* ont cherché à démontrer que les écoles font bien des différences et que l'action des enseignants et des directions d'établissement importe pour comprendre les différences de performance entre les classes ou les écoles d'un

système éducatif donné (Dumay, 2011). Ce courant s'est donc centré autour du thème de l'efficacité différenciée des établissements et des classes avec l'ambition initiale de fournir aux écoles des inputs leur permettant d'accroître leur capacité à faire progresser les élèves (Murphy et al., 1985).

Il existe différents facteurs d'efficacité des établissements sur la performance des élèves (Scheerens et Bosker, 1997). Le courant du *school mix* s'est attaché à explorer l'effet des caractéristiques propres aux écoles sous l'angle de la composition, c'est-à-dire «l'influence directe et indirecte exercée sur un élève par les autres élèves de son école» (Dumay et al., 2010, p.462). Selon Monseur et Crahay (2008), l'effet école est avant tout un effet d'agrégation sociale et académique: «plus le fonctionnement d'un système d'enseignement engendre des phénomènes d'agrégation sociale ou académique, plus il amplifie les différences entre élèves faibles et élèves forts et plus il amplifie ainsi les inégalités sociales devant l'école» (p. 60). D'après Thrupp (1999), l'effet de composition sociale et académique possède un caractère indirect médié par des processus internes à l'établissement. La présence d'une certaine composition au sein de l'école n'a pas d'effet en elle-même mais engendre des processus pédagogiques ou psychologiques spécifiques qui affectent les progressions académiques en modifiant la quantité et la qualité de l'instruction délivrée et les attitudes et comportements des élèves.

L'enquête COFO ne collecte pas d'informations sur les établissements fréquentés par les élèves. Dès lors, les possibilités de traitement de la problématique de l'effet établissement découlent uniquement de l'appariement des données avec celles extraites des bases de données scolaires cantonales. Les informations disponibles dans ces dernières, destinées à la gestion administrative de l'école, limitent les analyses aux seuls effets de composition décrits ci-dessus. Bien qu'il existe une littérature relativement abondante sur le sujet, étudier l'effet établissement sous cet angle présente un intérêt au-delà de ce que les résultats pourraient potentiellement apporter pour le monitoring des systèmes scolaires concernés. Produire de la connaissance sur l'effet établissement au Tessin et à Genève est essentiel puisque la recherche a déjà montré que ce dernier est tributaire des caractéristiques des systèmes éducatifs (Monseur et Crahay, 2008).

Questions de recherche et hypothèses

L'objectif principal de cet article est d'étudier l'existence de différences d'efficacité des établissements en matière d'acquisition des compétences de base en mathématiques et, le cas échéant, de voir si ces différences sont liées à la composition sociale et académique de la population scolarisée. Le traitement de ces questions nécessite la mise en œuvre d'une démarche analytique en trois étapes clairement explicitée par Duru-Bellat (2001). La première consiste à observer des *effets bruts* d'établissements, c'est-à-dire à s'assurer qu'il existe effectivement

des différences de résultats entre les différentes entités. En d'autres termes, il s'agit de déterminer quelle est, au sein de chaque canton, la part des différences de résultats au test COFO liée à l'établissement fréquenté. La seconde étape est l'identification d'*effets nets* définis comme «le *plus* ou le *moins* en termes de progression associé à la fréquentation d'un établissement pour des élèves ayant un niveau initial identique et des caractéristiques socio-démographiques comparables» (p.323). Dans quelle mesure les différences de résultats entre établissements subsistent-elles une fois contrôlées les caractéristiques sociodémographiques et scolaires des élèves qu'ils accueillent (Thrupp, 1998) ? Les premiers résultats de l'enquête COFO ont montré que l'origine sociale, le genre, le statut migratoire ou la langue parlée à la maison expliquent en partie les différences de chances d'atteindre les compétences de base (Consortium COFO, 2019). Mais les établissements performant-ils toujours diversement quand l'effet de ces caractéristiques individuelles est pris en compte ? Si tel est le cas, on tentera alors, dans la dernière étape de la démarche analytique, d'expliquer les *effets nets* en prenant en compte les caractéristiques des établissements, à savoir la composition sociale et académique du public accueilli.

Compte tenu des résultats qui figurent dans la littérature, on s'attend à observer des *effets bruts* d'une ampleur assez modeste puisque selon Elliott (1996) entre 8% et 15% de la variance des résultats des élèves s'expliquent par les différences entre écoles, la plupart des estimations se situant dans le bas de cette fourchette. D'autres sources indiquent que l'effet-école serait plutôt de l'ordre de 4 à 5% (Bianco et Bressoux, 2009). Par ailleurs il est assez probable que l'on parvienne également à mettre en évidence des *effets nets* car, dans bon nombre de travaux, les caractéristiques sociodémographiques et scolaires des élèves rendent compte d'une part importante mais non exhaustive des différences entre écoles. Enfin, il apparaît plus délicat de formuler une hypothèse sur un effet établissement correspondant, au moins partiellement, à un effet de composition sociale et académique du public accueilli car les résultats de recherche restent encore assez contrastés.

Données et méthode

Les résultats des élèves au test COFO constituent la variable principale d'intérêt. Conformément au modèle généralisé dans les enquêtes internationales (Vrignaud, 2006), ces résultats se présentent sous la forme de 20 valeurs plausibles fournissant une base de données qui permet d'estimer de façon non biaisée l'intervalle dans lequel se situe la compétence de groupes d'élèves (Von Davier et al., 2009). La particularité de COFO réside toutefois dans le fait que les 20 valeurs plausibles ont été dichotomisées en compétences fondamentales atteintes ou non (Angelone et Keller, 2019). Il s'agit là d'un point important car, au-delà des traditionnelles implications sur la manière de réaliser les analyses découlant d'une estimation de

l'atteinte des compétences par un ensemble de valeurs plausibles (Pham et al., 2019)¹, disposer de résultats exprimés sous la forme de deux modalités conduit à privilégier des modèles statistiques adaptés à ce type de variables dépendantes, notamment la régression logistique. Les caractéristiques des élèves (genre, statut migratoire, origine sociale et langue parlée à la maison) sont, elles aussi, collectées par le biais de l'enquête. Les analyses reposent sur des échantillons de 625 élèves de 11^e année HarmoS et 35 établissements pour le Tessin et de 612 élèves et 19 établissements pour Genève. Dans les deux cantons tous les établissements ont participé au test. Au sein de ces entités de 10 à 36 élèves ont été sélectionnés au Tessin et de 19 à 48 élèves à Genève. Les deux échantillons ont été pondérés pour être représentatifs de toute la population d'élèves de 11^e année HarmoS (Consortium COFO, 2019; Verner et Helbling, 2019).

Les données COFO ont dû être complétées par des informations extraites des bases cantonales. Cette mise en commun a été nécessaire pour caractériser la composition de la population d'un établissement. En effet, l'échantillon des seuls élèves qui ont participé au test, bien que représentatif des élèves de 11^e, ne permet pas de caractériser de façon intégrale la composition d'un établissement, ce qu'on peut en revanche faire grâce aux données cantonales qui recensent tous les élèves. Pour la composition académique, deux indicateurs ont été utilisés: les pourcentages d'élèves non promus et d'élèves scolarisés dans le programme le plus exigeant (deux niveaux A en mathématiques et en allemand au Tessin et filière à exigences élevées à Genève). D'autre part, selon Willms (1992) ou Lauder et Hughes (1990), certains effets de composition mis en évidence dans la littérature seraient, en réalité, des artifices dus à un contrôle lacunaire des variables caractérisant les élèves et tout particulièrement leur niveau initial. Les données COFO, comme celles issues de la plupart des évaluations à large échelle, ne donnent aucune indication de ce type mais le recours aux données cantonales permet de s'affranchir de cette limite. Ainsi, le contrôle du niveau initial des élèves soumis au test COFO a été effectué grâce à leur moyenne annuelle de fin de 10^e en mathématiques et au programme scolaire suivi lors de cette même année (i.e. les niveaux en mathématiques et en allemand pour le Tessin et les filières pour Genève).

Sur le plan méthodologique, les modèles logistiques multiniveaux (Sommet et Morselli, 2017) sont apparus comme les plus appropriés pour cette recherche puisqu'ils opèrent une décomposition de la variance totale des résultats au test en une variance individuelle et une variance contextuelle de façon simultanée. Ainsi, élèves et établissements sont considérés comme deux sources distinctes de variation de l'atteinte des compétences fondamentales. De façon plus précise, la mise en œuvre de ces modèles est tout à fait conforme à la procédure en trois étapes décrite par Duru-Bellat (2001). Dans un premier temps, on quantifie la part des différences de résultats liée à l'établissement à l'aide d'un modèle «vide», simple décomposition de la variance totale des résultats en une part de variance inter-établissements et une part de variance intra-établissements (*effets*

bruts). Dans une seconde étape, l'introduction des caractéristiques individuelles des élèves permet d'identifier des *effets nets*, c'est-à-dire de voir s'il subsiste une variance inter-établissements statistiquement significative une fois contrôlées des variables telles que le genre, l'origine sociale ou le niveau initial en mathématiques. Le cas échéant, en prenant en compte les caractéristiques des établissements, on pourra statuer sur la nature de l'effet en validant ou réfutant l'hypothèse d'un impact de la composition sociale et académique du public accueilli.

En pratique, la réalisation du modèle «vide» n'a posé aucune difficulté. En revanche, en raison de la taille trop limitée des échantillons cantonaux, les deux étapes suivantes n'ont pu être mises en œuvre de façon satisfaisante sur le plan statistique. La question de la taille des échantillons suffisante pour réaliser des modèles multiniveaux est complexe et reste encore un sujet de discussion. Selon Maas et Hox (2005), dans ces modèles la taille de l'échantillon au niveau des groupes est, en général, plus importante que la taille totale de l'échantillon où un grand nombre de sujets vient, en partie, compenser un faible nombre de groupes. Au Tessin et à Genève le nombre de groupes est limité puisque tous les établissements ont participé à l'enquête. Dès lors, c'est le nombre d'élèves échantillonnés au sein de chacun d'eux qui se révèle insuffisant. A titre de comparaison, ce type d'analyses a déjà été fait avec les données PISA pour ces cantons, mais la taille des échantillons était plutôt de l'ordre de 1'000 sujets que de 600².

L'impossibilité de mener à son terme la régression logistique multiniveau initialement envisagée a conduit à retenir une autre approche toujours basée sur la notion de valeur ajoutée mais davantage inspirée de ce qui a été développé en France pour analyser les performances des lycées au baccalauréat (Evain et Evrard, 2017). Cette approche fait une comparaison entre ce que devrait être la réussite des élèves d'un établissement compte tenu de la nature de son public (l'attendu), et la réussite réelle de ces mêmes élèves (l'observé). L'écart entre les deux mesures donne la «valeur ajoutée» (Felouzis, 2005). On évalue ainsi toujours l'action propre d'un établissement, ce qu'il a « ajouté au niveau initial des élèves qu'il a reçus » (Evain et Evrard, 2017, p.93) en éliminant l'incidence des facteurs de réussite scolaire qui lui sont extérieurs et en comparant la réussite de ses élèves à celle d'élèves semblables scolarisés dans des établissements comparables.

La première étape du calcul est l'estimation d'une régression logistique sur l'ensemble de l'échantillon de chaque canton. Elle vise à quantifier l'effet du genre, du statut migratoire, de la langue parlée à la maison, du statut socio-économique, du niveau initial et de la filière (uniquement pour Genève) sur les performances en mathématiques. Sur la base des coefficients significatifs, des probabilités théoriques d'atteinte des compétences fondamentales ont été estimées pour chaque élève. Lorsqu'elles sont supérieures ou égales à 50% on considère que l'élève atteint les compétences fondamentales. Cette valeur de césure de 0.5 est la plus couramment utilisée en régression logistique pour examiner dans quelle mesure un modèle permet de bien classer les sujets dans

leur groupe d'appartenance (Hosmer et Lemeshow, 2000). Un examen détaillé a montré que le nombre d'élèves dont les probabilités estimées sont proches de la valeur de césure est faible ce qui confère une certaine robustesse à la classification obtenue. Enfin, les résultats ont été agrégés au niveau des établissements pour calculer une proportion théorique d'élèves qui atteint les compétences fondamentales. Ce calcul a été fait seulement sur les établissements d'au moins 20 élèves. Le résultat correspond à une valeur attendue qui, comparée à la valeur réellement observée, donne la valeur ajoutée.

Enfin, dans une dernière étape plus descriptive, correspondant à la troisième phase de la démarche analytique détaillée plus haut, on cherchera à voir si les établissements ayant une valeur ajoutée significative se distinguent des autres du point de vue de leur composition sociodémographique et académique.

Résultats

Analyse logistique multiniveau: les apports du modèle «vide»

Le modèle «vide» met en évidence une variance inter-établissements modeste mais statistiquement significative qui indique que la probabilité d'atteindre les compétences fondamentales varie d'un établissement à l'autre. Au Tessin et à Genève respectivement 1.5% et 3% des différences de résultats sont liées à l'établissement fréquenté. Evidemment une partie de cette variance s'explique par le fait que les établissements accueillent des élèves dont les caractéristiques sont différentes. Il serait alors intéressant de voir comment cette variance évolue lorsqu'on contrôle certaines d'entre elles, mais la taille des échantillons empêche de réaliser cette opération de façon satisfaisante. Toutefois, le modèle vide atteste qu'une partie des chances d'atteindre les compétences fondamentales est bien liée à l'établissement fréquenté ce qui justifie la poursuite des analyses et légitime la mise en oeuvre de modèles complémentaires permettant le calcul d'une valeur ajoutée pour chaque établissement.

Valeur ajoutée: une majorité d'établissements réussit de façon conforme à ce que l'on peut attendre

Détaillons tout d'abord l'effet des différentes caractéristiques des élèves introduites dans les régressions logistiques utilisées pour le calcul de la valeur ajoutée. Le niveau initial en mathématiques (note annuelle de 10^e) possède un effet positif et statistiquement significatif sur la probabilité d'atteindre les compétences fondamentales au Tessin et à Genève (Figures 1 et 2). Par ailleurs, au Tessin le fait d'avoir été scolarisé dans le niveau le moins exigeant en mathématiques possède un effet négatif sur la probabilité de réussir au test. Dans une logique similaire, à Genève, être issu des filières à exigences moyennes ou élémentaires constitue, toutes choses égales par ailleurs, un facteur qui atténue les chances d'atteindre

les compétences fondamentales, tout comme le fait de fréquenter la filière à exigences moyennes au cours de l'année de passation des tests. En outre, au Tessin les filles possèdent moins de chances d'atteindre ces compétences. Enfin, dans les deux cantons, l'origine sociale, la langue parlée à la maison et le statut migratoire ne possèdent aucun effet une fois contrôlé le niveau initial.

Figure 1: Modèle logistique estimant la probabilité d'atteindre les compétences fondamentales en mathématiques dans le canton du Tessin.

	Canton du Tessin				
	Coefficients	Erreur-type	Probabilité critique	Significativité	Exp(b)
Constante	-0.821	1.117	0.461	ns	0.440
Niveau initial					
Note de mathématiques en 10 ^e	1.324	0.238	<0.001	**	3.760
Niveau de mathématiques en 10 ^e - Niveau B (réf. Niveau A)	-2.952	0.335	<0.001	**	0.052
Genre (réf. Garçon)	-0.594	0.307	0.055	ns	0.552
Statut migratoire (réf. Natif)					
Migrant 2 nd e génération	-0.512	0.471	0.279	ns	0.600
Migrant 1 ^{re} e génération	0.128	0.453	0.780	ns	1.137
Origine sociale					
Hisei08	0.005	0.007	0.510	ns	1.005
Langue parlée à la maison (réf. Uniquement langue de scolarisation)					
Plusieurs langues dont celle de scolarisation	0.190	0.338	0.576	ns	1.209
Uniquement autre(s) langue(s)	-0.242	0.611	0.690	ns	0.785

** : significatif à 0.01; * : significatif à 0.05; ns : non significatif.

Figure 2: Modèle logistique estimant la probabilité d'atteindre les compétences fondamentales en mathématiques dans le canton de Genève.

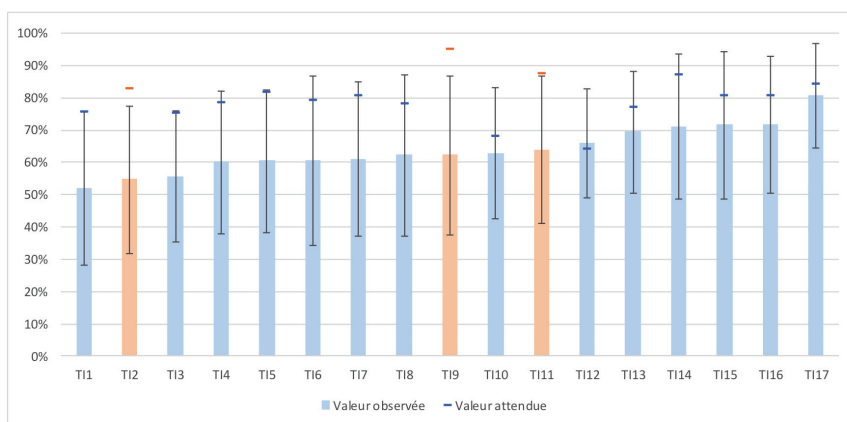
	Canton de Genève				
	Coefficients	Erreur-type	Probabilité critique	Significativité	Exp(b)
Constante	-3.127	1.323	0.020	*	0.044
Niveau initial					
Note de mathématiques en 10 ^e	1.283	0.295	<0.001	**	3.609
Filière 10 ^e à exigences moyennes (réf. Filière 10 ^e à exigences élevées)	-1.539	0.594	0.011	*	0.215
Filière 10 ^e à exigences élémentaires (réf. Filière 10 ^e à exigences élevées)	-3.667	0.988	<0.001	**	0.026
Filière de 11 ^e (réf. Filière à exigences élevées)					
Filière à exigences moyennes	-1.452	0.520	0.006	**	0.234
Filière à exigences élémentaires	-1.329	0.775	0.086	ns	0.265
Genre (réf. Garçon)	-0.662	0.371	0.077	ns	0.516
Statut migratoire (réf. Natif)					
Migrant 2 nd e génération	-0.425	0.420	0.314	ns	0.654
Migrant 1 ^{re} e génération	-0.191	0.565	0.734	ns	0.826
Origine sociale					
Hisei08	0.009	0.008	0.259	ns	1.009
Langue parlée à la maison (réf. Uniquement langue de scolarisation)					
Plusieurs langues dont celle de scolarisation	-0.242	0.404	0.549	ns	0.785
Uniquement autre(s) langue(s)	-0.021	0.681	0.976	ns	0.979

** : significatif à 0.01; * : significatif à 0.05; ns : non significatif.

Sur la base des deux modèles qui précèdent, il est possible d'estimer pour chaque établissement un taux d'atteinte des compétences fondamentales attendu. La comparaison avec la valeur observée révèle que la grande majorité des écoles obtient des résultats conformes à ce que l'on peut attendre. En effet, au Tessin on ne dénombre que trois établissements pour lesquels on peut conclure que les deux taux diffèrent significativement (Figures 3 et 4). Il s'agit d'établissements dont

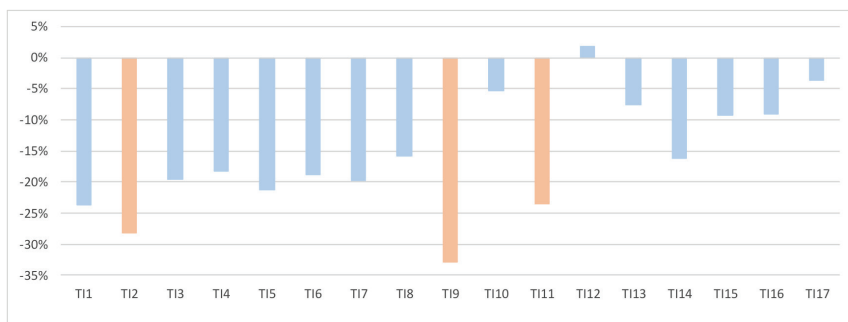
les résultats sont plus faibles que ce à quoi on pourrait s'attendre compte tenu de la population qu'ils scolarisent. A Genève, trois établissements se trouvent également dans la même situation (Figures 5 et 6). On notera qu'au Tessin les établissements concernés occupent des positions assez différentes puisque deux d'entre eux se situent plutôt dans le milieu de la distribution des résultats alors que le troisième fait partie des moins performants. A Genève, en revanche, tous les établissements concernés font partie des écoles les moins performantes du canton.

Figure 3: Taux d'atteinte des compétences fondamentales observés et attendus selon l'établissement dans le canton du Tessin



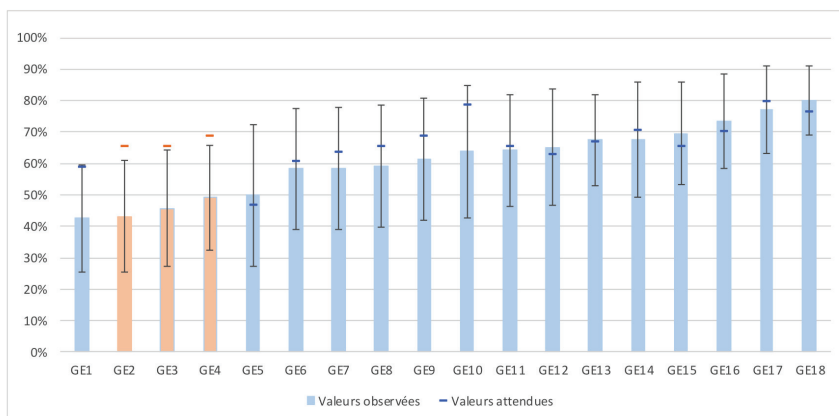
Note: les écoles sont triées par ordre croissant selon leur niveau d'atteinte des compétences fondamentales. Seules les écoles où plus de 20 élèves ont été testés sont représentées.

Figure 4: Valeur ajoutée des établissements dans le canton du Tessin



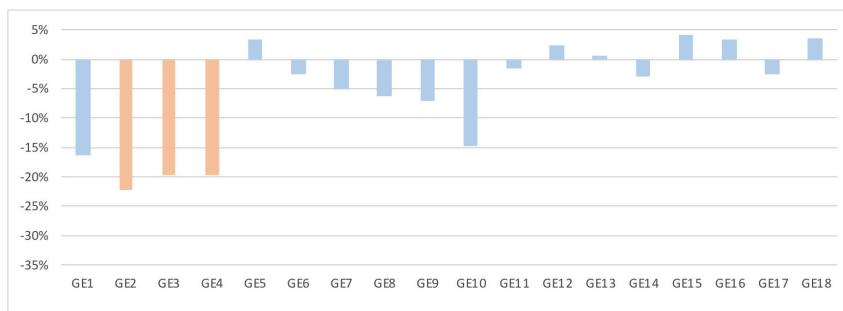
Note: les écoles sont triées par ordre croissant selon leur niveau d'atteinte des compétences fondamentales. Seules les barres en orange représentent des valeurs ajoutées statistiquement significatives.

Figure 5: Taux d'atteinte des compétences fondamentales observés et attendus selon l'établissement dans le canton de Genève



Note: les écoles sont triées par ordre croissant selon leur niveau d'atteinte des compétences fondamentales. Seules les écoles où plus de 20 élèves ont été testés sont représentées.

Figure 6: Valeur ajoutée des établissements dans le canton de Genève



Note: les écoles sont triées par ordre croissant selon leur niveau d'atteinte des compétences fondamentales. Seules les barres en orange représentent des valeurs ajoutées statistiquement significatives.

Un lien entre caractéristiques de composition des écoles et valeur ajoutée négative ?

Une dernière étape consiste à voir, de façon descriptive, si les six établissements possédant une valeur ajoutée négative ont des caractéristiques de composition qui les distinguent des valeurs moyennes prévalant dans leur canton. On considère que, sur chacune des caractéristiques de composition, la valeur d'un établissement diffère de la norme lorsqu'elle n'est pas comprise dans un intervalle défini par la valeur moyenne plus ou moins un écart-type³. Au Tessin, les trois établissements concernés ne se distinguent en rien des autres du point de vue de

leur composition. En revanche, la situation est un peu différente à Genève où deux des établissements concernés possèdent à la fois une proportion d'élèves scolarisés dans la filière à exigence élevée plus faible et une proportion d'élèves allophones supérieure à la moyenne. L'un d'eux possède aussi des proportions d'élèves non promus et d'élèves issus de milieu favorisé respectivement plus élevée et plus faible que les valeurs moyennes. Le dernier établissement concerné ne se distingue pas, quant à lui, de la norme en matière de composition de la population scolarisée. Ainsi, il semble bien difficile de tirer des conclusions évidentes sur un possible lien entre composition et efficacité des établissements.

Discussion et conclusions

Cet article avait pour but de voir dans quelle mesure COFO peut apporter à un niveau local des informations permettant de juger de l'action des établissements, en particulier de leur degré d'efficacité. L'évaluation de l'atteinte des compétences fondamentales est un instrument de monitoring et de régulation pour la Suisse mais aussi pour les cantons et fournir des informations sur ce qui se joue à l'intérieur de chacun d'eux est une nécessité. A ce titre, cet article constitue un premier essai d'exploration des possibilités d'approfondissement des analyses nationales sur la base des informations dont disposent les cantons (Consortium COFO, 2019).

Les analyses ont montré des différences modestes de résultats entre établissements: 1.5% au Tessin et 3% à Genève. Ces dernières apparaissent plus faibles que celles mentionnées généralement dans la littérature et en particulier pour les cantons concernés (Origoni, 2007; Petrucci et al., 2018). De plus, en raison de l'impossibilité de prendre en compte le niveau intermédiaire correspondant à la classe, on ne peut totalement écarter une surestimation de la variance inter-établissements (Opdenakker et Van Damme, 2000). Dès lors, il convient de faire preuve de prudence quant à l'existence d'un effet établissement. On peut d'ailleurs s'interroger sur les pistes qui expliquent la «quasi-absence» d'un tel effet alors même qu'il existe un consensus autour de l'idée que l'établissement scolaire est, en lui-même, une source d'inégalité des acquis des élèves (Dumay, 2004).

Il faut vraisemblablement revenir au type de compétences mesurées dans COFO pour mieux comprendre ces résultats. Ces tests visent à évaluer l'atteinte de compétences de base devant être maîtrisées par la totalité des élèves. Dès lors, ils sont par nature assez peu discriminants et génèrent peu de variance au niveau des élèves et des établissements. Comme le mentionne De Ketele (2009), « il faut pouvoir admettre que les conclusions tirées sont déterminées par la nature des performances évaluées» (p. 92). Une observation attentive des taux d'atteinte des compétences fondamentales des établissements et des intervalles de confiance associés montre qu'au Tessin et à Genève seuls quelques établissements situés aux extrémités de la distribution des résultats peuvent être considérés comme étant

différents du point de vue de leur réussite. Ceci explique probablement pourquoi une analyse réalisée sur l'ensemble d'un canton met en évidence une variance inter-établissements statistiquement significative mais pour autant relativement modeste. D'autre part, l'effet établissement est tributaire des caractéristiques des systèmes éducatifs (Monseur et Crahay, 2008). Au Tessin et à Genève les établissements disposent de peu d'autonomie et la plupart des principaux facteurs qui régissent leur fonctionnement sont fixés par voie réglementaire et s'appliquent de manière identique sur tout le territoire. Ceci atténue la variabilité entre établissements et limite les possibilités de voir émerger un effet important.

La deuxième partie de l'analyse confirme ce qui précède. La grande majorité des établissements obtiennent des résultats tout à fait conformes à ceux auxquels on peut s'attendre compte tenu de la population d'élèves qu'ils scolarisent. Il s'agit d'un constat plutôt positif en matière d'équité car il signifie que, la plupart du temps, la probabilité d'atteindre les compétences fondamentales pour un élève donné n'est pas impactée par l'établissement de scolarisation. Le fonctionnement des établissements ne semble pas accentuer les inégalités de départ liées aux caractéristiques sociodémographiques. Quelques établissements ont cependant des résultats observés plus faibles que ceux escomptés mais aucun lien entre leur sous-performance et leur composition n'a pu être établi.

De façon plus générale, cette recherche conduit à s'interroger sur la capacité réelle de COFO à fournir des informations sur ce qui se joue à l'intérieur des cantons. Tout d'abord, le type de compétences évalué est peu enclin à générer des différences et réduit les possibilités d'analyse. L'enquête COFO représente une mobilisation conséquente de ressources au niveau national – plus de 22'000 élèves concernés en 2016 – et constitue un outil indispensable au pilotage de la formation en Suisse, mais elle n'en demeure pas moins relativement limitée pour dépasser la comparaison «macro» entre cantons. La faible taille des échantillons et les informations collectées réduisent bien vite les opportunités d'analyses plus sophistiquées, en particulier celles sur les effets du contexte. Bien que cette enquête ait permis de comparer l'acquisition des compétences fondamentales dans l'ensemble des cantons, elle n'est cependant qu'un premier essai de mesure au niveau Suisse. Dès lors, les résultats de la présente recherche pourraient constituer une modeste contribution aux réflexions en cours sur le devenir et les évolutions souhaitables de cette évaluation à large échelle.

Notes

- 1 Le principe est que les analyses statistiques sont élaborées à partir de l'ensemble des valeurs plausibles et non d'une seule ou d'un agrégat de celles-ci. Les régressions logistiques ont été réalisées avec le package R BIFIEsurvey (BIFIE, 2018) et les modèles multiniveaux avec la version 7 du logiciel HLM.
- 2 Dans PISA 2012 l'échantillon tessinois comportait 1'081 élèves scolarisés dans 35 établissements (de 25 à 35 élèves selon l'établissement). Pour Genève, on dénombrait 920 élèves scolarisés dans 19 établissements (de 31 à 69 élèves selon l'établissement) (Petrucci et al., 2018).

- 3 Les valeurs moyennes ont été calculées à partir des seuls établissements pour lesquels on dispose d'un échantillon d'au moins 20 élèves ayant participé au test COFO (établissements représentés sur les Figures 4 et 6).

Références bibliographiques

- Angelone, D., et Keller, F. (2019). *ÜGK 2016 Mathematik. Technische Dokumentation zu Testentwicklung und Skalierung*. Aarau: Geschäftsstelle der Aufgabendatenbank EDK (ADB).
- Bianco, M., et Bressoux, P. (2009). Effet-classe et effet-maître dans l'enseignement primaire: vers un enseignement efficace de la compréhension ? Dans X. Dumay et V. Dupriez (dir.), *L'efficacité dans l'enseignement. Promesses et zones d'ombre* (p. 35-54). De Boeck.
- BIFIE (2018, 28 décembre). BIFIEsurvey: Tools for survey statistics in educational assessment. R package version 2.5-44. <https://CRAN.R-project.org/package=BIFIEsurvey>.
- Bressoux, P. (2006) Effet-maître et pratiques de classe. Dans E. Bourgeois et G. Chapelle (dir.), *Apprendre et faire apprendre* (p. 214-226). PUF.
- Bressoux, P. (1994). Note de synthèse. Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*, 108, 91-137.
- CDIP (2007). *Accord intercantonal sur l'harmonisation de la scolarité obligatoire (concordat HarmoS) du 14 juin 2007*. CDIP. <https://edudoc.ch/record/24710?ln=fr>
- CDIP (2003). *Mesures consécutives à PISA 2000: plan d'action, décision de la CDIP du 12 juin 2003*. http://www.edudoc.ch/static/web/arbeiten/pisa2000_aktplan_f.pdf
- Coleman, J.S., Campbell, E., Hobson, C., Mcpartland, J., Mood, A., Weinfeld, F., et York, R. (1966). *Equality of Educational Opportunity*. Government printing office.
- Consortium COFO (dir.) (2019). *Vérification de l'atteinte des compétences fondamentales. Rapport national COFO 2016: mathématiques 11e année scolaire*. CDIP et SRED. <https://doi.org/10.18747/PHSGcoll3/id/384>
- Consortium PISA.ch. (2011). *PISA 2009 Risultati regionali e cantonali*. UFFT/CDPE e Consorzio PISA.ch
- De Ketele, J.M. (2009). Avancées et limites des recherches sur l'efficacité. Dans X. Dumay et V. Dupriez (dir.), *L'efficacité dans l'enseignement. Promesses et zones d'ombre* (p. 89-99). De Boeck.
- Dumay, X. (2011). La recherche sur l'efficacité de l'enseignement ; transformations politiques et modèle d'évaluation. *Les dossiers des sciences de l'éducation*, 25, 57-68.
- Dumay, X. (2004). Effet établissement : effet de composition et/ou effet des pratiques managériales et pédagogiques ? Un état du débat. *Les cahiers de recherche en éducation et formation*, 34.
- Dumay, X., et Dupriez, V. (2009). *L'efficacité dans l'enseignement: Promesses et zones d'ombre*. De Boeck Supérieur. doi:10.3917/dbu.dumay.2009.01.
- Dumay, X., Dupriez, V., et Maroy, C. (2010). Ségrégation entre écoles, effets de la composition scolaire et inégalités de résultats. *Revue française de sociologie*, 51(3), 461-480. doi: 10.3917/rfs.513.0461.
- Duru-Bellat, M. (2001). Effets maîtres, effets établissements : quelle responsabilité pour l'école ? *Schweizerische Zeitschrift für Bildungswissenschaften*, 23(2), 321-337.
- Elliott, J. (1996). School effectiveness research and its critics: alternative visions of schooling. *Cambridge Journal of Education*, 26(2), 199-224.
- Evain, F., et Evrard, L. (2017). Une meilleure mesure de la performance des lycées : refonte de la méthodologie des IVAL (session 2015). *Education et formations, Ministère de l'éducation nationale, Mathématiques : clefs de lecture des résultats TIMSS 2015*, 91-116.
- Felouzis, G. (2005). Performances et "valeur ajoutée" des lycées: le marché scolaire fait des différences. *Revue française de sociologie*, 46, 3-36.
- Gorard, S. (2006). Is there a school mix effect? *Educational review*, 58(1), 87-94.

- Grisay, A. (2006). Chapitre 16. Que savons-nous de l'«effet établissement»? Dans G. Chapelle (dir.), *Améliorer l'école* (p. 215-230). Presses Universitaires de France.
- Hosmer, D.W., et Lemeshow, S. (2000) *Applied logistic regression* (2^e éd.). John Wiley & Sons, Inc.
- Lauder, H., et Hughes, D. (1990). Social inequalities and differences in school outcome. *New Zealand Journal of Educational Studies*, 25, 37-60.
- Loeb, S., et Figlio, D. (2011). School accountability. Dans E. A. Hanushek, S. Machin et L. Woessmann (dir.), *Handbook of the Economics of Education*, 3 (p. 383-423). North Holland.
- Maas, C.J., et Hox, J.J. (2005). Sufficient sample size for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1(3), 86-92.
- Monseur, C., et Crahay, M. (2008). Composition académique et sociale des établissements, efficacité et inégalités scolaires: une comparaison internationale. *Revue française de pédagogie*, 164, 55-66.
- Murphy, J., Weil, W., Hallinger, P., et Mitman, A. (1985). School effectiveness: a conceptual framework. *The Educational Forum*, 49(3), 361-374.
- Opendakker, M.-C., et Van Damme, J. (2000). The Importance of Identifying Levels in Multilevel Analysis: An Illustration of the Effects of Ignoring the Top or Intermediate Levels in School Effectiveness Research. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 11(1), 103-130.
- Origoni, P. (2007). *Equi non per caso. I risultati dell'indagine PISA 2003 in Ticino*. Ricerche in educazione. Ufficio studio e ricerche.
- Pham, G., Helbling, L., Verner, M., Petrucci, F., Angelone, D., et Ambrosetti, A. (2019). *ÜGK – COFO – VeCoF 2016 results: Technical appendices*. Pädagogische Hochschule St. Gallen (PHSG) et Service de la Recherche en éducation (SRED).
- Petrucci, F., Ambrosetti, A., Fenaroli, S., et Egloff, M. (2018). Effet établissements sur la réussite des élèves au Tessin et à Genève. Service de la recherche en éducation et Centro innovazione e ricerca sui sistemi educativi.
- Rutkowski, L., Von Davier, M., et Rutkowski, D. (2014). A brief introduction to modern international large-scale assessment. Dans L. Rutkowski, M. Von Davier et D. Rutkowski (dir.), *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis*. Taylor and Francis.
- Sheerens, J., et Bosker, R.J. (1997). *The foundations of educational effectiveness*. Pergamon.
- Sommet, N., et Morselli, D. (2017). Keep calm and learn multilevel logistic modeling: a simplified three-step procedure using Stata, R, Mplus and SPSS. *International Review of Social Psychology*, 30(1), 203-218.
- Thrupp, M. (1999). *Schools making a difference – Let's be realistic!*. McGraw-Hill Education.
- Thrupp, M. (1998). The art of the possible: organizing and managing high and low socioeconomic schools. *Journal of Education Policy*, 13(2), 197-219.
- Verner, M., et Helbling, L. (2019). *Sampling ÜGK 2016. Technischer Bericht zu Stichproben-design, Gewichtung und Varianzschätzung bei der Überprüfung des Erreichens der Grundkompetenzen 2016*. Institut für Bildungsevaluation, assoziiertes Institut der Universität Zürich.
- Von Davier, M., Gonzalez, E., et Mislevy, R.J. (2009). What are plausible values and why are they useful? Dans D. Hastedt et M. von Davier (dir.), *IERI Monograph Series. Issues and Methodologies in Large-Scale Assesments* (IERI Monograph Series, vol. 2, p. 9-36)
- Vrignaud, P. (2006). La mesure de la littéracie dans PISA : la méthodologie est la réponse, mais quelle était la question ?. *Revue française de pédagogie*, 157, 27-41.
- Willms, J.D. (1992). *Monitoring school performance*. Falmer Press.

Mots-clés: Effet du contexte, valeur ajoutée, efficacité de l'école, composition de la population, Tessin et Genève, atteinte des compétences fondamentales (COFO)

Die Bedeutung der Schulen für die Schülerinnen- und Schülerleistungen. Eine Analyse der Resultate der ÜGK-Erhebung 2016

Zusammenfassung

Die Leistungen der Schülerinnen und Schüler hängen aus Sicht der Schulleffektivitätsforschung sowohl von individuellen Schülermerkmalen als auch vom schulischen Kontext ab. Inwieweit kann die breit angelegte Erhebung zur „Überprüfung der Grundkompetenzen“ (ÜGK) auf Schweizer Ebene den einzelnen Kantonen Informationen zur Wirksamkeit ihrer Schulen liefern? Im Tessin und in Genf erscheint der Anteil der Schülerleistungen, der auf Unterschiede zwischen Schulen zurückzuführen ist, gering zu sein. Eine Berechnung des Zugewinns (value added) zeigt, dass die Mehrheit der Schulen Ergebnisse erzielt, die im Bereich des Erwartbaren liegen. Nur einige wenige Schulen zeigen schlechtere Ergebnisse als erwartet; es zeigte sich jedoch kein Zusammenhang zwischen geringeren Leistungen und der Zusammensetzung der Schülerpopulation.

Schlagerworte: Kontext-Effekt, Zugewinn (value-added), Schulleistung, Zusammensetzung der Schülerpopulation, Tessin und Genf, Überprüfung der Grundkompetenzen (ÜGK)

Impatto degli istituti scolastici sulle performance degli allievi. Uno studio basato sui risultati dell'indagine VeCoF 2016

Riassunto

Secondo la corrente dello *school effectiveness*, i risultati scolastici degli allievi dipendono sia dalle loro caratteristiche individuali sia dalle caratteristiche del contesto scolastico in cui tali allievi sono inseriti. In che misura l'indagine VeCoF, una valutazione su larga scala effettuata a livello svizzero, può fornire ai cantoni informazioni per rilevare l'efficacia delle proprie scuole? In Ticino e a Ginevra, le differenze di risultati scolastici tra le scuole appaiono estremamente modeste. Il calcolo del valore aggiunto mostra che la grande maggioranza degli istituti ottiene risultati in linea con quanto ci si può attendere. Solo in pochi istituti si rilevano risultati inferiori alle aspettative, ma non è stato trovato alcun nesso tra le scarse prestazioni scolastiche individuate in tali istituti e le caratteristiche della relativa popolazione studentesca.

Parole chiave: Effetto del contesto, valore aggiunto, efficacia scuole, composizione della popolazione, Ticino e Ginevra, verifica delle competenze fondamentali (VeCoF)

Impact of schools on student performance. A study based on the ÜGK 2016 survey

Summary

According to the school effectiveness research stream, students' performance depends on their individual characteristics and their schooling context. To what extent can the ÜGK survey, a large-scale assessment carried out at the Swiss level, provide the cantons with information for judging their schools' effectiveness? In Ticino and Geneva, the proportion of differences in results linked to the school appears to be extremely modest. Value added calculations show that the vast majority of schools achieve results in line with those that can be expected. Only a few have lower than expected results, but no link between their underperformance and their students' population demographic can be established.

Keywords: Contextual effect, added value, school effectiveness, population demographic, Ticino and Geneva, achievement of core competencies (ÜGK).

Franck Petrucci, collaborateur de recherche au Service de la recherche en éducation du canton de Genève (SRED) et chercheur associé au Centro innovazione e ricerca sui sistemi educativi (CIRSE – SUPSI). Contribue à l'exploitation des données et à la réalisation des rapports nationaux COFO. Principaux domaines de recherche: effet du contexte scolaire sur les performances des élèves. Service de la recherche en éducation du canton de Genève (SRED), Quai du Rhône 12, CH-1205 Genève
E-Mail: franck.petrucci@etat.ge.ch

Alice Ambrosetti, ricercatrice del Centro competenze innovazione e ricerca sui sistemi educativi (CIRSE) della Scuola universitaria professionale della Svizzera italiana (SUPSI). Attiva principalmente nell'implementazione della Verifica delle competenze fondamentali (VeCoF) per la Svizzera italiana. Principali aree di ricerca: indagini su larga scala sulle competenze scolastiche. Scuola universitaria professionale della Svizzera italiana (SUPSI), Dipartimento formazione e apprendimento, Piazza San Francesco 19, CH-6600 Locarno
E-Mail: alice.ambrosetti@supsi.ch

Francesca Crotta, ricercatrice del Centro competenze innovazione e ricerca sui sistemi educativi (CIRSE) della Scuola universitaria professionale della Svizzera italiana (SUPSI). Attiva principalmente nell'implementazione della Verifica delle competenze fondamentali (VeCoF) per la Svizzera italiana. Principali aree di ricerca: indagini su larga scala sulle competenze scolastiche. Scuola universitaria professionale della Svizzera italiana (SUPSI), Dipartimento formazione e apprendimento, Piazza San Francesco 19, CH-6600 Locarno
E-Mail: francesca.crotta@supsi.ch

Christian Nidegger, jusqu'en décembre 2019, collaborateur de recherche au Service de la recherche en éducation du canton de Genève (SRED). Responsable de la vérification de l'atteinte des compétences fondamentales (COFO) pour la Suisse romande. Principaux domaines de recherche: évaluation des compétences des élèves dans les enquêtes à large échelle. Service de la recherche en éducation du canton de Genève (SRED), Quai du Rhône 12, CH-1205 Genève
E-Mail: christian.nidegger@etat.ge.ch