



HAL
open science

**Systèmes d'accountability basés sur la performance:
types, logiques instrumentales et effets sur l'efficacité et
l'équité scolaires des systèmes éducatifs d'Europe et du
Canada. Une étude comparative à partir des données
PISA 2012**

Annelise Voisin

► **To cite this version:**

Annelise Voisin. Systèmes d'accountability basés sur la performance: types, logiques instrumentales et effets sur l'efficacité et l'équité scolaires des systèmes éducatifs d'Europe et du Canada. Une étude comparative à partir des données PISA 2012 . Education. Université de Montréal, 2017. Français. NNT: . tel-01705972

HAL Id: tel-01705972

<https://shs.hal.science/tel-01705972>

Submitted on 15 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Montréal

THÈSE

En vue de l'obtention du grade de Philosophiæ Doctor (PHD) en éducation comparée et
fondements de l'éducation

**Systemes d'*accountability* basés sur la performance: types, logiques
instrumentales et effets sur l'efficacité et l'équité scolaires des systèmes
éducatifs d'Europe et du Canada.
Une étude comparative à partir des données PISA 2012.**

Présentée et soutenue publiquement par

ANNELISE VOISIN

Le 3 novembre 2017

Directeur de thèse :

Christian Maroy, Professeur, Université de Montréal

Examineur externe :

Pierre Doray, Professeur, Université du Québec à Montréal

Président rapporteur :

Nathalie Loye, Professeur, Université de Montréal

Autre membre du Jury :

Canisius Kamanzi, Professeur, Université de Montréal

© Annelise Voisin, 2017



RÉSUMÉ

L'*accountability* est devenue, dans les dernières décennies, la voie privilégiée d'une gouvernance renouvelée des systèmes d'éducation. Sous l'influence du *New Public Management*, l'introduction de systèmes d'*accountability* basés sur la performance (SABP) est envisagée comme une solution politique à la crise postulée des systèmes éducatifs. Les SABP devraient ainsi concourir à l'amélioration des performances des élèves et à la réduction des inégalités scolaires, néanmoins la littérature peine à voir émerger un consensus sur ces questions.

Privilégiant une acception large, mais opérationnelle de l'*accountability*, prenant en compte les instruments mobilisés au sein des systèmes éducatifs pour réguler le fonctionnement des établissements scolaires sur base de leurs résultats, cette recherche quantitative comparative internationale analyse les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada.

L'étude repose sur une analyse secondaire des données du PISA 2012 et porte sur un échantillon de 59 systèmes d'éducation. Elle mobilise des analyses multi-variées (analyse en composantes principales, analyses de classification hiérarchique, ANOVA, régressions simples et multiples) pour d'une part caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada et, d'autre part, analyser leurs effets sur l'efficacité et l'équité scolaires.

Les résultats du premier volet de recherche font émerger trois logiques instrumentales au cœur des SABP : 1) une logique d'évaluation et de régulation selon la performance ; 2) une logique reposant sur la mise en place d'incitatifs directs pour les enseignants, 3) une logique articulant comparaison externe et dispositifs internes d'amélioration des performances des écoles. Soutenues par des instruments distincts, ces logiques s'articulent au sein de quatre types de SABP qui permettent de différencier les systèmes d'éducation d'Europe et du Canada : 1) une *accountability* de faible intensité et peu instrumentée, 2) une *accountability* axée principalement sur l'évaluation de la performance, 3) une *accountability* articulant suivi de la performance et comparaison externe/amélioration interne de la performance des établissements scolaires; 4) une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants.

Les résultats du second volet de recherche suggèrent que les effets des SABP sur l'efficacité et d'équité scolaire des systèmes éducatifs de l'échantillon sont, le cas échéant, faibles et s'annulent le plus souvent lorsque des facteurs clefs sont pris en compte (dépenses publiques en éducation, variation du niveau socioéconomique des élèves, stratification horizontale et verticale des systèmes scolaires). Si la mise en place d'incitatifs pour le corps enseignants, et le type de SABP privilégiant ce type de dispositifs, apparaissent significativement liés à des degrés d'efficacité scolaire plus faibles, d'inégalités de résultats et des acquis de base plus élevés, nos résultats invitent finalement à dresser un constat plus que mitigé quant à la capacité des SABP à tenir leurs promesses d'efficacité et d'équité scolaires.

Mots-clefs : *Accountability*, typologie empirique, efficacité et équité scolaires, PISA 2012, étude quantitative internationale, éducation comparée, systèmes éducatifs Europe et Canada.

ABSTRACT

In the last decades accountability, has become a centerpiece of the renewed governance of the education systems. Under the influence of *New Public Management*, the introduction of performance-based accountability systems (PBAS) is considered a political solution to the supposed crisis affecting education systems. While PBAS are aimed at improving student performance and at closing the achievement gap, there is little evidence that PBAS have a positive impact on school effectiveness and equity in education systems.

This international comparative quantitative study aims to provide some insights into the relationship between PBAS, effectiveness and equity (equal opportunities of results, equality of outcomes and equality of basic skills) in the European and Canadian education systems. It uses a broad but operational acceptance of the concept of accountability that takes into account the result-based regulation tools that are put in place at the school level in order to steer the system.

The study is based on a secondary analysis of OECD PISA 2012 data. It is based on a sample of 59 education systems and use multivariate analyses (principal component analysis, cluster analysis, ANOVA, simple and multiple regressions) to characterize the PBAS in the education systems of Europe and Canada on one side, and to analyse their impact on effectiveness and equity in these education systems on the other.

Three instrumental logics emerge from the results, which underpin PBAS: 1) a logic that rests on monitoring and evaluating the actors's performance at the school level; 2) a logic that relies on direct incentives for teachers; 3) a logic that rests on school performance comparisons and involves internal mechanisms aimed at improving school results. Underpinned by distinct regulation tools, these logics are articulated within empirical types of PBAS which allow us to differentiate between the education systems of Europe and Canada: 1) an accountability of lesser intensity, 2) an accountability focused mainly on evaluation and performance monitoring, 3) an accountability articulating performance monitoring, external comparison/internal improvement of school performance, 4) a high intensity accountability favoring incentives for teachers.

Nevertheless, we find little evidence that PBAS have a significant influence on effectiveness and equity in education systems. The effects of PBAS, if any, are in most cases weak and are cancelled out when other key factors are taken into account (public spending in education, variation in student socioeconomic backgrounds, horizontal and vertical stratification of education systems). However, incentives for teachers and PBAS that favor this type of device seem to be significantly linked to lower effectiveness levels, and to higher inequalities in outcomes and basic skill levels. In the end, our results suggest that PBAS do not provide the effectiveness and equity in education systems they promise.

Keywords: Accountability systems, empirical typology, effectiveness and equity in education, PISA 2012, international quantitative study, comparative education, Europe and Canada.

TABLE DES MATIÈRES

INTRODUCTION	1
PREMIÈRE PARTIE : PROBLÉMATIQUE ET CADRAGES	7
CHAPITRE 1 : PROBLÉMATIQUE	8
SECTION 1 : Une nouvelle régulation de l'éducation	8
1. 1 Convergences partielles des politiques éducatives et régulation par les résultats	9
1.2. L' <i>Accountability</i> en éducation : caractériser l'objet	12
1.2.1 De l' <i>accountability</i>	12
1.2.2 L' <i>accountability</i> comme composante de la « bonne gouvernance ».....	13
1.2.3 L' <i>accountability</i> comme mécanisme	14
1.2.4 Formes organisationnelles traditionnelles de l' <i>accountability</i>	16
1.3 Avènement de la régulation de l'éducation par les résultats et « nouvelle <i>accountability</i> »	18
1.4 Systèmes d' <i>accountability</i> basés sur la performance, instruments privilégiés	20
1.4.1 Standards.....	23
1.4.2 Dispositifs d'évaluation	25
1.4.3 Reddition de comptes et conséquences associées	26
1.5 Stratégies de changement.....	28
1.6 L'influence du <i>New Public Management</i>	29
1.7 Une rhétorique commune.....	30
1.8 Conclusion	34
SECTION 2 : Effets des systèmes d' <i>accountability</i> basés sur la performance sur l'efficacité et l'équité scolaires, revue de la littérature	37
1.9 De l'efficacité.....	38
1.9.1 Effets de l'introduction du <i>testing</i> sur l'efficacité scolaire, les études comparatives inter-États (États-Unis)	38
1.9.2 Effets des SABP à forts et faibles enjeux, les études comparatives inter-États.....	41
1.9.3 Le cas texan.....	43
1.9.4 Effets des SABP, les études comparatives internationales	44
1.10 De l'équité.....	45
1.11 Effets imprévus	48
1.12 Conclusion	51
1.13 Synthèse de la problématique et perspectives générales de recherche	52

CHAPITRE 2 : CADRAGE THÉORIQUE.....	57
SECTION 1 : Régulation de l'éducation et systèmes d' <i>accountability</i> basés sur la performance	58
2.1 La régulation des systèmes d'éducation, les apports de la sociologie	58
2.2 Modes de régulation institutionnelle et instruments d'action publique	63
2.3 Retour sur la notion d' <i>accountability</i>	64
2.4 Caractériser les SABP par leurs instruments	66
2.5 Conclusion	70
SECTION 2 : Efficacité et équité scolaires	72
2.6 De l'efficacité.....	72
2.7 De l'équité et des inégalités	76
2.7.1 Égalité des chances et méritocratie	78
2.7.2 Égalité de résultats et égalité des acquis de base	82
2.8 Conclusion	85
2.9 Synthèse et démarche générale de recherche.....	85
CHAPITRE 3 : METHODOLOGIE.....	91
3.1 Une étude comparative quantitative internationale.....	91
3.2 Les données.....	94
3.2.1 L'enquête PISA 2012.....	94
3.2.2 Les bases de données internationales, la littérature institutionnelle et empirique	98
3.3 L'échantillon.....	99
3.4 Démarche méthodologique et méthodes d'analyse des données : les deux volets empiriques de la recherche.....	100
3.4.1 Premier volet empirique de recherche : démarche méthodologique et méthodes d'analyses des données	102
3.4.1.1 Postulats et objectifs	102
3.4.1.2 Opérationnalisation du concept de systèmes d' <i>accountability</i> basés sur la performance et variables mobilisées.....	103
3.4.1.3 Méthodes d'analyse des données	107
3.4.1.3.1 Analyse en composantes principales	108
3.4.1.3.2 Analyse typologique : analyse de classification hiérarchique	111
3.4.2 Second volet empirique de recherche : démarche méthodologique et méthodes d'analyses des données	115
3.4.2.1 Hypothèses et objectifs	116
3.4.2.2 Opérationnalisation des concepts d'efficacité et d'équité scolaires : indicateurs et variables mobilisées.....	116
3.4.2.2.1 Degré d'efficacité scolaire	117
3.4.2.2.2 Degré d'équité scolaire	118
3.4.2.2.3 Les variables de contrôle	121

3.4.2.4 Les méthodes d'analyse des données.....	123
3.4.2.4.1 Analyses des effets des composantes instrumentales des SABP sur l'efficacité et l'équité scolaires	123
3.4.2.4.2 Analyses des effets des types empiriques de SABP sur l'efficacité et l'équité scolaires.....	126
3.4.2.5 Considérations méthodologiques	130
3.5 Conclusion	131
DEUXIÈME PARTIE : RÉSULTATS DE RECHERCHE.....	133
CHAPITRE 4 : LA DIVERSITÉ DES SYSTÈMES D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE EN EUROPE ET AU CANADA – UNE APPROCHE EMPIRIQUE PAR LES INSTRUMENTS/MODES DE RÉGULATION	134
SECTION 1 : Composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance	136
4.1 Résultats des analyses en composantes principales	136
4.1.1 Première composante instrumentale : évaluation et régulation selon la performance.....	139
4.1.2 Deuxième composante instrumentale : enjeux et incitatifs directs pour les enseignants	140
4.1.3 Troisième composante instrumentale : comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires	142
4.2 Une comparaison des systèmes éducatifs d'Europe et du Canada selon les différentes composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance	144
4.3 Synthèse et interprétation des composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance	147
SECTION 2 : types empiriques différenciés de systèmes d' <i>accountability</i> basés sur la performance	153
4.4 Résultats des analyses de classification hiérarchique	153
4.4.1 Répartition de l'échantillon au sein des types empiriques.....	154
4.4.2 Caractérisation des types empiriques à partir de leurs moyennes sur les composantes	158
4.4.2.1 Premier type empirique : système d' <i>accountability</i> de faible intensité et peu instrumentée.....	159
4.4.2.2 Deuxième type empirique : <i>accountability</i> axée sur l'évaluation de la performance	160
4.4.2.3 Troisième type empirique : <i>accountability</i> axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance	161
4.4.2.4 Quatrième type empirique : <i>accountability</i> de forte intensité prévoyant des incitatifs directs pour les enseignants	161
4.5 Interprétation des résultats autour des types de systèmes d' <i>accountability</i> basés sur la performance	162
4.5.1 Type 1 : <i>accountability</i> de faible intensité et peu instrumentée.....	162
4.5.2 Type 2 : <i>accountability</i> axée sur l'évaluation de la performance	164

4.5.3 Type 3 : <i>accountability</i> axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance	166
4.5.4 Type 4: <i>accountability</i> de forte intensité prévoyant des incitatifs directs pour les enseignants	168
4.6 Conclusion	171
CHAPITRE 5 : QUELS SONT LES EFFETS DES SYSTÈMES D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE SUR L'EFFICACITÉ ET L'ÉQUITÉ SCOLAIRES ?	173
SECTION 1 : Composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance, efficacité et équité scolaires des systèmes éducatifs d'Europe et du Canada : quelles relations, quels effets ?	175
5.1 Analyse descriptive comparative des degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada	175
5.2 Résultats des analyses de corrélations	181
5.2.1 Relations (corrélations) entre composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance, efficacité et équité scolaires	182
5.2.1.1 Efficacité et composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance	182
5.2.1.2 Inégalité de résultats et composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance	182
5.2.1.3 Inégalités des chances de résultats et composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance.....	183
5.2.1.4 Inégalités des acquis de base et composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance.....	183
5.2.2 Relations (corrélations) entre degrés d'efficacité et d'équité scolaires et variables de contrôle	184
5.2.3 Relations (corrélations) entre composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance et variables de contrôle	185
5.3 Effets bruts des composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance: résultats des analyses de régression simple	186
5.4 Effets nets des composantes instrumentales des systèmes d' <i>accountability</i> basés sur la performance sur les degrés d'efficacité et d'équité scolaires : résultats des analyses de régressions multiples.....	187
5.4.1 Les relations entre variables de contrôle, degrés d'efficacité et d'équité scolaires	190
5.4.2 Effets nets de la composante instrumentale des incitatifs directs articulés aux évaluations des enseignants : résultats des analyses de régressions multiples	191
5.4.3 Effets nets de la composante de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires : résultats des analyses de régressions multiples.....	192
5.5 Synthèse	193
SECTION 2 : Types de systèmes d' <i>accountability</i> basés sur la performance, degrés d'efficacité et d'équité scolaires, des relations différenciées?	197

5.6 Descriptif des degrés d'efficacité et d'équité scolaires par types de systèmes d' <i>accountability</i> basés sur la performance	198
5.7 Résultats des ANOVA	202
5.8 Effets nets des types de systèmes d' <i>accountability</i> basés sur la performance sur les degrés d'efficacité et d'équité scolaires : résultats des analyses de régressions multiples	204
5.8.1 Types de systèmes d' <i>accountability</i> basés sur la performance et efficacité scolaire	205
5.8.2 Types de systèmes d' <i>accountability</i> basés sur la performance et degrés d'inégalités de résultats	206
5.8.3 Types de systèmes d' <i>accountability</i> basés sur la performance et degrés d'inégalités des chances de résultats	206
5.8.4 Types de systèmes d' <i>accountability</i> basés sur la performance et degrés d'inégalités des acquis de bases.....	207
5.9 Synthèse	208
5.10 Synthèse et conclusion des résultats de recherche	210
CHAPITRE 6 : DISCUSSION GÉNÉRALE	214
6.1. Discussion du premier volet de recherche : caractériser les systèmes d' <i>accountability</i> basés sur la performance dans les systèmes éducatifs d'Europe et du Canada	214
6.1.1 Des instruments aux logiques instrumentales des systèmes d' <i>accountability</i> basés sur la performance	215
6.1.2 Les logiques instrumentales comme base d'une typologie exploratoire des systèmes d' <i>accountability</i> basés sur la performance.....	217
6.1.3 Logiques et types empiriques de systèmes d' <i>accountability</i> basés sur la performance : la question des enjeux pour les acteurs	219
6.2 Discussion du second volet de recherche : effets des systèmes d' <i>accountability</i> basés sur la performance sur l'efficacité et l'équité scolaires, un constat mitigé	221
6.2.1 Systèmes d' <i>accountability</i> basés sur la performance, efficacité et équité : des effets qui demeurent faibles et instables	221
6.2.2 Relation entre systèmes d' <i>accountability</i> basés sur la performance et réduction des inégalités scolaires d'origine socioéconomique : des effets inexistantes	224
6.3 Stratification des systèmes éducatifs et dépenses publiques en éducation : freins ou vecteurs d'équité scolaire et d'efficacité scolaire	225
6.4 La rhétorique politique à l'épreuve de résultats factuels	227
6.5 Des effets normatifs et systémiques qui méritent d'être questionnés	228
6.5.1 La rationalisation du pilotage des systèmes éducatifs : des liens aux systèmes de production	228
6.5.2 Une rationalisation par la mesure	230
6.6 Systèmes d' <i>accountability</i> basés sur la performance et principes de justice : des conceptions sous-jacentes à expliciter	231
6.6.1 Justice scolaire : entre idéal d'égalité des chances et idéal d'égalité de résultats.....	231

6.6.2 Pour une vision plus ambitieuse de la justice scolaire.....	233
6.7 Une autre voie possible?	234
6.7.1 La centration sur les « produits » : vers des objectifs plus larges.....	234
6.7.2 Enjeux et responsabilisation des acteurs : vers une responsabilité collective	235
CONCLUSION.....	237
LIMITES ET PERSPECTIVES DE RECHERCHE.....	240
BIBLIOGRAPHIE.....	243

LISTE DES TABLEAUX

Tableau 1 : Tableau synoptique du design de recherche	90
Tableau 2 : Sources de données mobilisées	99
Tableau 3 : Tableau synoptique du design de recherche	101
Tableau 4 : Tableau synoptique du volet empirique 1 de recherche.....	102
Tableau 5 : Opérationnalisation du concept de système d' <i>accountability</i> basés sur la performance, variables mobilisées (PISA 2012)	105
Tableau 6 : Opérationnalisation du concept de système d' <i>accountability</i> basé sur la performance, méthodes d'analyses des données et variables mobilisées	114
Tableau 7 : Tableau synoptique du volet empirique 2 de recherche.....	115
Tableau 8 : Variables mobilisées dans les analyses de la thèse	123
Tableau 9 : Variables intégrées aux modèles de régressions multiples	128
Tableau 10 : Tableau synoptique du volet 1 de recherche.....	135
Tableau 11 : Rappel, variables d'intérêt, instruments des SABP	137
Tableau 12 : Résultats de l'ACP, indice de Kaiser-Meyer-Olkin et test de Bartlett	138
Tableau 13 : Indices de saturation des variables sur composante d'évaluation selon la performance	139
Tableau 14 : Indices de saturation des variables de la composante incitatifs et enjeux pour le corps enseignant.....	141
Tableau 15 : Indices de saturation des variables de la composante de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires	143
Tableau 16 : Première ébauche de la typologie empirique des SABP.....	157
Tableau 17 : Caractéristiques des types empiriques de SABP : moyenne et écart-type pour chaque composante instrumentale et différences avec les autres types empiriques.	159
Tableau 18 : Rappel, tableau synoptique du volet empirique 2 de recherche.....	174
Tableau 19 : Relations (corrélations) entre composantes des SABP, efficacité et équité scolaires, variables de contrôle	181
Tableau 20 : Synthèse des relations (corrélations) entre composantes instrumentales des SABP, efficacité et équité scolaires	184
Tableau 21 : Effets bruts des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires — résultats des analyses de régression simple	187
Tableau 22 : Effets nets des composantes instrumentales sur le degré d'efficacité et d'équité scolaires — résultats des analyses de régressions multiples.....	189
Tableau 23 : Types empiriques de SABP, efficacité et équité scolaires : synthèse des résultats des ANOVA	203
Tableau 24 : Effets des types empiriques de SABP sur les degrés d'efficacité et d'équité scolaires	205

LISTE DES FIGURES

Figure 1 : <i>Accountability</i> comme mécanisme (Bovens, 2010, p. 454)	15
Figure 2 : Schéma conceptuel — régulation des systèmes éducatifs par les résultats et systèmes d' <i>accountability</i> basés sur la performance (SABP)	71
Figure 3 : Schéma conceptuel systèmes d' <i>accountability</i> basés sur la performance (SABP), efficacité et équité scolaires	87
Figure 4 : Composante instrumentale d'évaluation et de régulation selon la performance, comparaison des systèmes éducatifs de l'échantillon	145
Figure 5 : Composante instrumentale des enjeux et incitatifs directs pour les enseignants, comparaison des systèmes éducatifs de l'échantillon	146
Figure 6 : Composante instrumentale de comparaison externe et dispositifs internes d'amélioration des performances de l'école	147
Figure 7 : Dendrogramme illustrant la répartition de l'échantillon au sein des types empiriques de SABP (méthode de Ward comme algorithme de regroupement et distance euclidienne au carré)	154
Figure 8 : Degré d'efficacité scolaire (mathématiques) – comparatif des systèmes éducatifs de l'échantillon	176
Figure 9 : Degré d'inégalité de résultats (mathématiques) – comparatif des systèmes éducatifs de l'échantillon	177
Figure 10 : Degré d'inégalité des chances de résultats (mathématiques) – comparatif des systèmes éducatifs de l'échantillon	178
Figure 11 : Degré d'inégalité des acquis de base (mathématiques) – comparatif des systèmes éducatifs de l'échantillon	178
Figure 12 : Efficacité et inégalités des chances de résultats, profils croisés des systèmes éducatifs de l'échantillon	180
Figure 13 : Efficacité scolaire et types de SABP – examen des distributions	198
Figure 14 : Inégalités de résultats et types de SABP – examen des distributions	199
Figure 15 : Inégalités des chances de résultats et types de SABP – examen des distributions ..	200
Figure 16 : Inégalités des acquis de base et types de SABP – examen des distributions	201

ANNEXES

ANNEXE 1	258
ANNEXE 2	262
ANNEXE 3	264
ANNEXE 4	265
ANNEXE 5	269
ANNEXE 6	278
ANNEXE 7	281
ANNEXE 8	282

LISTE DES SIGLES, ABRÉVIATIONS ET ACRONYMES

ACP : Analyse en Composantes Principales

ACH : Analyse de classification hiérarchique (cluster)

ANOVA : Analyse de variance

BEL CFB&GER : Belgique communauté francophone et germanophone

ECOCE : *End of Course Exams*

ET : Écart-type

HSEEE : *High-School Examination Exit Exams* (examens de sortie du secondaire)

KMO : Kayser-Meyer-Olkin

M : Moyenne

MCT : *Minimum Competency Testing*

NAEP : *National Assessment of Educational Progress*

NCLB : *No Child Left Behind*

OCDE : Organisme de Coopération et de Développement Économique

PIB : Produit Intérieur Brut

PISA : *Program International for Student Assessment*

SABP : Système(s) d'*accountability* basés sur la performance

SBR : *Standard Based Reform*

SPSS : *Statistic Package for Social Sciences*

UNESCO : Organisation des Nations Unies pour l'éducation, la Science et la Culture

À ma mère, mon plus solide garde fou, mon plus grand soutien.

À ma grand-mère, mon ciel et mon refuge.

À mon frère, pour ce qu'il est et représente à mes yeux.

À mon père et ses conseils durant cette aventure.

À Michel Guintini et Cécile Sarda.

REMERCIEMENTS

Mes sincères remerciements vont en premier lieu à Christian Maroy. En tant que directeur de thèse, pour la rigueur et la qualité de son suivi, pour son implication dans la supervision de cette thèse doctorale. Pour m'avoir offert les conditions qui ont rendu ce travail possible, pour avoir respecté et soutenu mes choix tout au long de ce parcours. Mes remerciements s'adressent aussi à Christian Maroy en tant que directeur de la Chaire de Recherche du Canada en politiques éducatives. Pour nous avoir offert la possibilité de faire nos armes en tant que jeunes chercheurs dans un environnement de travail exigeant et stimulant. Pour nous avoir associés activement aux travaux de recherche menés par la Chaire, pour les apprentissages que j'ai pu y réaliser et la formation de haut vol qui m'a été dispensée.

Je tiens à exprimer mes sincères remerciements aux membres du jury, Nathalie Loye (Université de Montréal), Canisius Kamanzi (Université de Montréal), Pierre Doray (Université du Québec à Montréal), pour avoir accepté d'évaluer ce travail de recherche ainsi que pour les conseils prodigués dans certaines étapes de sa réalisation.

Je remercie chaleureusement Carla Barroso Da Costa pour ses conseils avisés ainsi que pour sa disponibilité pour répondre à mes nombreuses questions quant aux analyses réalisées dans ce travail.

Un immense merci à Javier Nunez, Carla Barroso, Sarah Mainich, à mon père, à ma mère, à Roxane pour leurs relectures de tout ou partie du manuscrit ; à Tàmhas Ó Móráin, Zoe Wake Hyde et Mathieu Jackson pour leur aide de dernière minute.

Mes remerciements vont par ailleurs aux professeurs du département qui ont participé à ma formation doctorale, mais aussi à tous les professeurs que j'ai pu solliciter dans différentes phases de la réalisation de cette thèse.

Je remercie Élodie Bellarbre, pour sa disponibilité et son aide lors de mon séjour de recherche à l'IREDU.

Merci à Sylvie Cartier avec qui j'ai eu la chance de travailler dans les premières années.

Merci à Claudine Jomphe et François Bowen du Vice-décanat aux études supérieures et à la recherche de la Faculté des Sciences de l'Éducation, aux membres du département dont Lucie Lefrançois, Chantal Viau, Paule Campeau, à Jean-François Durnin d'EPC Bio, toujours disponibles pour répondre à nos questions.

Mes remerciements vont par ailleurs aux collègues de la Chaire de Recherche du Canada en politiques éducatives, ceux de la première heure, Cécile Mathou et Samuel Vaillancourt, ceux arrivés plus tard dont Michel Diatta, Mathieu Jackson, Gonzague Yerly et à tous les membres de l'équipe. Car ils m'ont chacun aidée, d'une manière ou d'une autre, à faire avancer ce projet dans une ambiance amicale (voire festive !). Ce fut pour moi un réel plaisir que de partager ces années avec une si belle équipe.

Je n'oublie pas les collègues que j'ai eu la chance de rencontrer au sein de la Faculté et ceux avec qui j'ai eu l'occasion de travailler, spécialement Éliane Mourad, Pierre-Marc Leblanc et Julie Arseneault.

J'ai ici une pensée toute particulière pour Jorge Araujo Lozano et Elham Darchinian, pour leur générosité de cœur et les pensées toujours positives qu'ils m'ont témoignées.

Je remercie de tout cœur mes très chères amies Sarah et Carla, pour leur soutien tout au long du chemin, pour leur appui dans mon parcours professionnel ainsi que dans ma vie personnelle; pour leur bienveillance, leurs conseils, pour les joies partagées et pour avoir traversé ensemble les bons moments, comme les mauvais. Merci les filles.

Merci à Yuri qui a suivi mes pérégrinations hors des murs de l'université, m'a souvent permis de voir les choses sous un autre angle et a été un solide soutien ces dernières années.

Je remercie mon père pour son écoute attentive et ses conseils tout au long de ce parcours.

Merci à ma mère, femme formidable qui m'accompagne dans chacun des moments de ma vie, pour laquelle j'ai une immense gratitude, un profond respect et une grande admiration. Ces quelques lignes sont bien peu pour t'exprimer ma reconnaissance.

Finalement, mes remerciements vont à Javier. Pour ses conseils avisés, sa présence intelligente, sa patience et sa bonne humeur à toute épreuve, pour ses encouragements et son précieux soutien. Pour m'avoir redonné l'énergie, la sérénité et la confiance nécessaires pour finir ce travail. Parce que grâce à toi la dernière ligne droite a été une fête bien plus qu'un chemin de croix.

INTRODUCTION

Pièce maîtresse des réformes à l'œuvre dans le champ éducatif, l'*accountability* est devenue, dans les dernières décennies, la voie privilégiée d'une gouvernance renouvelée des systèmes d'éducation. Dans un mouvement de transformation des modes de coordination mis en place par les États, sous l'influence des principes du *New Public Management*, des systèmes d'*accountability* basés sur la performance (SABP) sont mis en œuvre dans la majorité des systèmes éducatifs des pays du globe pour réguler leur fonctionnement : aux États-Unis (Linn, 2000; Figlio & Loeb, 2011), au Chili (Fallabella, 2013), au Canada (Maroy, Mathou, Vaillancourt & Voisin, 2013), en Angleterre (Ozga, 2009), en Nouvelle-Zélande, ou encore dans certains pays d'Afrique comme le Kenya (Bruns, Filmer, & Patrinos, 2011).

Implémentés selon des modalités différenciées en fonction des particularités de contextes, les SABP reposent toutefois sur un ensemble d'instruments centraux tels que des standards, des dispositifs d'évaluation standardisés, des mécanismes de reddition de comptes auxquels sont attachées diverses conséquences pour les acteurs du champ éducatif. Il s'agit de gouverner par les nombres (Ozga, 2009 ; Felouzis, 2011), de mesurer et de suivre les résultats du système à la lumière d'objectifs quantifiés (des élèves, des écoles, du système éducatif dans son ensemble), de responsabiliser les acteurs (enseignants, écoles, instances de régulation intermédiaires) quant aux performances des élèves et de mettre cette responsabilisation au travail par le biais de mécanismes de reddition de comptes.

L'introduction des SABP au sein des systèmes éducatifs s'appuie ainsi sur une rhétorique commune, celle de l'efficacité, de la modernité, de la qualité. Dans un contexte de préoccupations grandissantes autour de l'efficacité et de l'équité des systèmes d'éducation (Hutmacher, Cochrane & Bottani, 2001), il est attendu des SABP qu'ils participent de l'amélioration du niveau de performances des élèves et qu'ils réduisent les inégalités scolaires persistantes au sein des systèmes éducatifs. Néanmoins, malgré la popularité des SABP auprès des décideurs politiques, un consensus empiriquement fondé peine à émerger sur ces questions.

Une littérature récente, mais prolifique s'est en effet intéressée aux effets des SABP et aux implications de ce que certains nomment la « nouvelle régulation de l'éducation par les résultats » dans le champ éducatif (Broadfoot, 2000, Ozga, 2009, Maroy, 2006, 2008).

D'un côté, des études quantitatives, principalement menées dans le contexte des États-Unis, interrogent les effets des SABP sur l'efficacité et l'équité des systèmes éducatifs (Carnoy & Loeb, 2002 ; Hanushek & Raymond, 2005 ; Harris & Herrington, 2006 ; Chiang, 2009 ; Ladd & Lauen, 2010, ect.). Elles mettent en exergue les variations qui existent dans les instruments au cœur des SABP, considérant ces derniers au travers de leur technicité et de leurs effets sur les performances scolaires.

D'autres travaux, dans le champ de la sociologie, analysent l'évolution des modes de régulation des systèmes d'éducation dans une perspective davantage critique. Ils mettent en exergue les implications normatives des changements qui traversent le champ éducatif et soulèvent des éléments de controverse quant à l'avènement de la « nouvelle régulation de l'éducation » (Demailly, 2001, 2005 ; Ball, 2003 ; Maroy, 2009 ; Ozga, 2009, etc.).

Divers constats émergent dès lors de ces travaux :

- Si nous laissons de côté les divergences des conclusions des études menées dans le contexte des États-Unis, il apparaît que les SABP auraient des effets positifs sur l'amélioration des performances des élèves. Toutefois ces effets s'observeraient principalement dans le cadre des SABP prévoyant des conséquences matérielles et sociales importantes pour les acteurs, dans certaines matières et pour les niveaux de scolarité les plus élevés (Lee, 2008).
- À l'inverse, les SABP tendraient à produire des effets négatifs en termes d'équité scolaire. Ils participeraient de l'augmentation de l'écart de réussite entre certains groupes d'élèves et de la ségrégation scolaire (Hanushek & Raymond, 2005), des taux de décrochage et de redoublement (Haney, 2000 cité Carnoy & Loeb, 2002).
- De plus, il apparaît que les instruments qui constituent le socle des SABP varient fortement selon les contextes où ils sont mis en œuvre. Les SABP revêtent dès lors des formes différenciées et leurs effets, le cas échéant, seraient fortement liés aux dispositifs sur lesquels ces derniers reposent.

- Cependant, peu d'études quantitatives comparatives se sont attelées à analyser les effets de différents types de SABP sur l'efficacité et l'équité scolaires hors des frontières des États-Unis.
- Finalement, les changements à l'œuvre au sein des systèmes d'éducation, dont la mise en œuvre des SABP est un des piliers centraux, s'inscrivent dans des tendances lourdes (Lessard, 2002) et participent de l'émergence d'un nouveau référent normatif en éducation (Demailly, 2001 ; Lessard & Meirieu, 2004).

Ces éléments, qui seront discutés dans notre thèse, permettent de poser le contexte de notre recherche doctorale et d'appuyer sa pertinence sociale et scientifique. D'un point de vue social, la présente recherche vise à produire des évidences empiriques permettant de répondre à la question centrale suivante : les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ? D'un point de vue scientifique, elle constitue une des rares études quantitatives comparatives internationales interrogeant les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation d'Europe et du Canada.

Si l'*accountability* est un concept ombrelle qui ne connaît pas de traduction littérale en français (Maroy & Voisin, 2017), nous privilégions une acception large de la notion, prenant en compte les instruments mobilisés au sein des systèmes éducatifs pour réguler le fonctionnement des établissements scolaires sur base de leurs résultats, pour analyser les effets des SABP sur l'efficacité et l'équité scolaires.

L'étude quantitative comparative au cœur de notre thèse doctorale repose sur une analyse secondaire des données du PISA 2012 et porte sur un échantillon de 59 systèmes d'éducation. Elle mobilise des analyses multi-variées (analyses en composantes principales, analyses de classification hiérarchique, ANOVA, régressions simples et multiples) pour caractériser d'une part les SABP dans les systèmes éducatifs d'Europe et du Canada et, d'autre part, interroger leurs effets sur l'efficacité et l'équité scolaires de ces systèmes d'éducation. Au-delà des résultats empiriques produits, la thèse interroge par ailleurs les tendances actuelles qui traversent le champ éducatif.

La thèse est constituée de 2 parties regroupant 6 chapitres. Les trois premiers chapitres, regroupés au sein d'une première partie, sont dédiés aux éléments de problématique ainsi qu'à la présentation des cadrages théoriques et méthodologiques. Le chapitre 1 prend comme ancrage

terminologique la notion d'*accountability* pour situer l'avènement de la « nouvelle régulation de l'éducation », et le développement des systèmes d'*accountability* basés sur la performance (SABP), dans un contexte de transformations des systèmes éducatifs. Nous y proposons par ailleurs une synthèse de littérature sur les effets des SABP sur l'efficacité et l'équité scolaires et présentons les perspectives générales de recherche. Le chapitre 2 pose le cadrage théorique. Nous y situons tout d'abord les SABP au travers des concepts de régulation et de modes de régulation institutionnelle, puis stabilisons les acceptions des notions d'efficacité et d'équité scolaires. Nous y précisons par ailleurs la démarche générale de recherche. Finalement, le cadrage méthodologique est posé au chapitre 3. Nous y présentons les éléments de méthodologie au cœur de notre étude quantitative comparative internationale : les données mobilisées issues de l'enquête PISA 2012, l'opérationnalisation des concepts centraux de la thèse, les postulats et hypothèses au cœur des volets empiriques déployés ainsi que les analyses réalisées.

Les chapitres 4 et 5 sont regroupés au sein d'une deuxième partie consacrée aux résultats de recherche. Le chapitre 4 présente les résultats de la première série d'analyses menées pour caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada (analyses en composantes principales, analyses de classification hiérarchiques). Le chapitre 5 présente les résultats de la deuxième série d'analyses permettant d'appréhender les effets des SABP sur les degrés d'efficacité et d'équité scolaires de ces systèmes d'éducation (corrélations, ANOVA, analyses de régressions simples et multiples).

Le chapitre 6 est consacré à la discussion générale de notre travail de thèse doctorale. Il synthétise et discute les principaux résultats de recherche. Nous y proposons par ailleurs une réflexion plus large, qui prolonge nos résultats de recherche, où nous formulons des pistes en termes davantage normatifs pour « penser » les SABP et les transformations à l'œuvre dans le champ éducatif.

Finalement, la conclusion met en exergue les apports de la thèse, ses limites et ouvre des perspectives de recherche futures.

What matters most is how well you walk through the fire.

Charles Bukowski

Les politiques d'éducation et de formation sont construites (...) dans des situations qui portent la marque de l'hétérogénéité et de la complexité caractéristiques de toute société, mais plus encore des sociétés contemporaines. (...) Construire une politique d'éducation c'est donc poser des priorités, choisir, mais aussi parfois exclure et presque toujours passer par des compromis.

Bernard Charlot

PREMIÈRE PARTIE : PROBLÉMATIQUE ET CADRAGES

CHAPITRE 1 : PROBLÉMATIQUE

Ce chapitre pose les éléments de problématique au cœur de notre recherche doctorale. Il est divisé en deux sections. Dans une première section, nous discutons de la notion d'*accountability* pour situer l'avènement de la « nouvelle régulation de l'éducation », et le développement des systèmes d'*accountability* basés sur la performance (SABP), dans un contexte de réformes des systèmes éducatifs. Nous discutons les instruments qui constituent le socle des SABP, des facteurs d'influence et des enjeux qui entourent leur mise en œuvre dans le champ éducatif.

La deuxième section présente et discute les conclusions des principales études, principalement états-uniennes, qui se sont intéressées aux effets des SABP sur l'efficacité et l'équité scolaires. Cette section permet de mettre en exergue les évidences, les limites, mais aussi les zones d'ombres que cette littérature suggère. À l'issue de cette section, nous mettons en exergue la pertinence de notre étude et présentons les perspectives générales de notre recherche.

SECTION 1 : Une nouvelle régulation de l'éducation

Cette section met en perspective l'avènement de la « nouvelle régulation de l'éducation » (Broadfoot, 2000; Maroy, 2008; Ozga, 2009) et le développement des systèmes d'*accountability* basés sur la performance dans une perspective internationale. Elle repose sur une analyse de la littérature et vise à mettre en relief différents points de convergence qui permettent de situer plus avant les transformations ayant cours au sein des systèmes d'éducation.

Nous y discutons tout d'abord des convergences partielles des politiques éducatives qui portent et participent de l'avènement d'une « nouvelle régulation » de l'éducation. Afin de situer notre propos dans le vaste univers sémantique qui entoure l'étude de la « nouvelle régulation des systèmes d'éducation », nous mobilisons la notion d'*accountability* comme point de départ à la problématisation de notre objet d'étude. Nous mettons en exergue les acceptions multiples du terme pour stabiliser la terminologie employée dans notre recherche doctorale. Après avoir présenté les différentes formes organisationnelles relativement traditionnelles de l'*accountability* dans le champ éducatif, nous proposons des éléments d'analyse permettant de saisir ce que certains conceptualisent comme le passage vers une « nouvelle *accountability* » (Carnoy & Loeb, 2002; Maroy & Voisin, 2013, 2014). En centrant notre propos sur les systèmes d'*accountability*

basés sur la performance (SABP), nous discutons par ailleurs les instruments au cœur de ces derniers en mettant l'accent, d'une part, sur leurs traits communs, d'autre part en faisant référence au cas de certains systèmes éducatifs pour introduire leur diversité. Ensuite, après avoir discuté la stratégie de changement commune au SABP, nous nous tournons vers les cadres du *New Public Management* et l'influence certaine de cette doctrine sur les changements en cours dans les trente dernières années. Finalement, nous mettons en exergue la rhétorique commune qui entoure le développement de l'*accountability* et des SABP dans le champ éducatif.

1. 1 Convergences partielles des politiques éducatives et régulation par les résultats

Si la massification et la démocratisation de l'accès à l'enseignement intervenues dans la deuxième moitié du XX^e siècle, se sont faites sous le joug d'un État providence, alors État éducateur principalement en charge du financement de l'éducation comme de la gestion des structures éducatives et des contenus scolaires, le rôle prépondérant de ce dernier va peu à peu évoluer et être remis en question à la fin des années 1970. Sur fond de crise économique, un cycle de politiques éducatives s'enclenche à l'échelle mondiale, porté d'une part par les demandes économiques et sociales qui émergent au sein des économies et des nations, d'autre part, appuyées par les grands opérateurs internationaux.

Bien que les réformes des systèmes scolaires portent sur des objets divers (curriculum, structures administratives, évaluation, financement de l'éducation), des tendances s'observent néanmoins dans « la façon dont l'État et son administration exercent la coordination et le contrôle sur l'ensemble du système » (Maroy, 2006, p. 281).

Si l'on s'intéresse en effet plus particulièrement à l'évolution des modes de régulation institutionnelle, entendus comme la « configuration spécifique des arrangements institutionnels effectivement prégnants dans un système éducatif » (Maroy, 2008, p. 5), des convergences partielles se dessinent, portées par les politiques éducatives et appuyées par des textes législatifs importants (*Education Reform Act* en Angleterre en 1988, la *Loi d'Orientation* de 1989 en France, l'*Educate America Act* en 1994 aux États-Unis, le *Décret mission* de 1997 en Communauté française de Belgique). Ces réformes s'articulent autour de la recherche d'un équilibre entre centralisation et décentralisation, de l'autonomie accrue des établissements, de la promotion et de l'assouplissement du choix de l'école, de l'évaluation externe des établissements et du système scolaire, de la diversification de l'offre scolaire et de l'augmentation du contrôle

du travail enseignant, enfin de l'usage de l'évaluation comme instrument de pilotage (Barroso, 2000; Duru-Bellat & Meuret, 2001; Maroy, 2006, 2008; Vandenberghe, 2001).

C'est dans ce contexte d'intense activité politique (Lessard, 2002) que la nouvelle régulation des systèmes d'éducation, autrement nommée régulation par les résultats, se développe depuis une vingtaine d'années dans la majeure partie des systèmes éducatifs des pays de l'OCDE.

L'Ontario entame dans les années 1990 une profonde réforme de son système éducatif. La Commission Royale de l'apprentissage est commanditée en 1993 en réponse à une inquiétude publique croissante sur la qualité de l'éducation. Elle pointe les incohérences du système éducatif ontarien et émet de nombreuses recommandations qui seront mises en œuvre dans une visée d'amélioration de la qualité du système, dans une logique de responsabilisation et d'imputabilité : centralisation et normalisation des curriculums, définition de standards et d'indicateurs de performance, mise en place d'évaluation standardisée à grande échelle, publication des résultats des élèves, mécanismes de reddition de comptes et de contrôle de la qualité du système.

Au Québec, la politique « prendre le virage du succès » (MELS, 1997) fait suite à une vaste réflexion sur l'éducation et engage un ensemble de mesures ayant trait à la refonte et à la normalisation des curriculums, à une nouvelle répartition des pouvoirs en éducation (Lessard & Desjardins, 2011). Dans la décennie 2000, deux lois majeures (loi 124 de 2002 et loi 88 de 2008) formalisent la mise en œuvre de la politique de Gestion Axée sur les Résultats (Maroy, Mathou, Vaillancourt, & Voisin, 2013). Ces lois engagent un mouvement de recentralisation du système éducatif. Elles instrumentent une stratégie d'alignement des différents « paliers » du système (ministère, instances de régulation intermédiaires, établissements scolaires) dans une logique *top-down* (Brassard, 2009). Cette stratégie repose sur la fixation d'objectifs chiffrés aux différents niveaux du système (central, intermédiaire, local), la mise en œuvre d'instruments de contractualisation permettant de formaliser des mécanismes de reddition de comptes sur les résultats.

Aux États-Unis, l'*Elementary and Secondary Education Act.* de 2002 et la politique *No Child Left Behind* (NCLB - Encadré 1) marquent un tournant majeur dans l'implantation d'un système d'imputabilité à forts enjeux orchestré à l'échelle fédérale. Il s'appuie sur la fixation de standards

de performances, l'évaluation et le *monitoring* des résultats des États, des districts et des établissements, et sur un système d'incitatifs et de sanctions à l'égard des agents éducatifs.

L'Angleterre, dans les années 2000, met elle aussi en place un système d'imputabilité à forts enjeux (Broadfoot, 2000). À la suite de la politique de décentralisation, de dévolution et de dérégulation menée dans l'ensemble des services publics anglais par le gouvernement Thatcher, l'État met en place un curriculum national (via l'*Education Reform Act* en 1988). Il définit par ailleurs un ensemble d'indicateurs de performance et centralise le contrôle des résultats des élèves et des établissements scolaires en s'appuyant sur le déploiement d'un appareil de production et de recueil de données sur le fonctionnement du système éducatif relativement sophistiqué (Ozga, 2009).

En France, la Loi Organique relative aux Lois de Finances de 2001 institue un dispositif de régulation « reposant sur le contrôle du budget » (Dutercq & Cuculou, 2013). Il s'applique à l'ensemble des secteurs publics et engage, dans le champ éducatif, la définition d'objectifs de performance par l'État central dans le cadre de projets « dévolus » aux chefs d'établissements. Si le pilotage du système éducatif par objectifs est présent en France depuis les années 1980 (Dutercq & Cuculou, 2013), différents dispositifs législatifs, dont le « triptyque instrumental contrat-plan-évaluation » (Pons, 2014, p. 59), renforcent la régulation de l'éducation par les résultats qui s'intensifie depuis le milieu des années 2000 et l'arrivée au pouvoir du gouvernement Sarkozy.

Les politiques qui prennent le nom de politiques d'*accountability* dans l'univers anglo-saxon, de pilotage par les résultats en Suisse, en France ou en Belgique francophone (Maroy, 2013) connaissent aujourd'hui une expansion internationale. Elles sont mises en œuvre dans la plupart des pays du globe bien que prenant des formes variées selon les contextes.

S'intéresser à la nouvelle régulation de l'éducation et à l'*accountability* dans le champ éducatif nécessite toutefois d'en clarifier les acceptions et les enjeux. Force est de constater, en effet, le vaste univers sémantique et conceptuel qui entoure l'étude de la « nouvelle régulation » et de l'*accountability* en éducation. Gouvernance par les résultats, pilotage par les résultats, reddition de comptes, redevabilité, imputabilité, *accountability* sont des termes en effet employés tantôt comme des synonymes référant à une même « réalité », tantôt comme des notions distinctes dont l'acception est, dans nombre de productions scientifiques, peu explicitée.

La sous-section suivante vise donc à stabiliser la terminologie employée dans notre recherche doctorale et à caractériser notre objet d'étude.

1.2. L'Accountability en éducation : caractériser l'objet

Nous avons fait le choix de convoquer la notion d'*accountability* comme point de départ à la problématisation de notre objet de recherche. Puisqu'elle ne trouve pas de traduction littérale en langue française qui permette de rendre compte de la complexité de l'objet, il nous a paru fructueux de mobiliser cette notion « importée » de la littérature anglo-saxonne pour en saisir les éléments constitutifs et en énoncer les enjeux dans le champ éducatif. C'est en outre au sein de la littérature anglo-saxonne que les travaux sur la question sont les plus nombreux, bien que ces derniers se développent aussi dans le monde francophone. L'« exemple » anglo-saxon d'*accountability* suscite l'intérêt de la communauté scientifique en même temps que les politiques du même nom tendent à se développer dans les quatre coins du globe.

Les sous-sections suivantes visent donc à cercler l'objet en discutant les acceptions courantes de la notion et les formes organisationnelles traditionnelles que revêt l'*accountability* en éducation.

1.2.1 De l'*accountability*

Accountability is one of those golden concepts that no one can be against. It is increasingly used in political discourse and policy documents because it conveys an image of transparency and trustworthiness. However, its evocative powers make it also a very elusive concept because it can mean many different things to different people, as anyone studying accountability will soon discover (Bovens, 2007, p. 447).

L'*accountability* est un concept complexe, polysémique qui recouvre de nombreuses acceptions dans la littérature. Bovens (2006, 2007, 2010) est l'un des rares auteurs qui, à notre connaissance, ait cherché à retracer la genèse de la notion. S'appuyant sur les travaux de Dubnick (2002 cité dans Bovens, 2007, p. 448) il met en exergue l'origine anglo-normande du terme qui réfère historiquement, et d'un point de vue sémantique, à un exercice comptable et financier, en d'autres termes à la traduction littérale du terme anglo-saxon *accounting* : comptabilité, finance¹.

¹ Après la conquête de l'Angleterre par les Normands en 1066, Guillaume le Conquérant (William 1^{er}) exige l'inventaire des biens possédés par tous les propriétaires terriens consignés dans les *Domesday Books*. Permettant de recenser les possessions du royaume, base comptable au prélèvement de l'impôt, les *Domesday Books* deviennent un

De manière concomitante à l'avènement des régimes démocratiques, l'*accountability* s'est progressivement distanciée de cette acception essentiellement liée à la finance où elle a donc une longue tradition. Elle devient une notion politique qui réfère à la responsabilité et à l'obligation démocratique des gouvernements élus de rendre comptes aux citoyens (Maroy & Voisin, 2017a). Parallèlement à l'influence des principes du *New Public Management* (Nouvelle Gestion Publique – NPM), cette notion prend par ailleurs un essor particulier dans le discours scientifique et des politiques publiques depuis les années 1980. Un des termes les plus usités dans l'administration publique (Mulgan, 2000 ; Sinclair, 1995 cités dans Veselý, 2013), elle recouvre des sens multiples et réfère à des « réalités » non moins diversifiées.

1.2.2 L'*accountability* comme composante de la « bonne gouvernance »

Largement convoqué dans le discours politique contemporain, la notion d'*accountability* y est mobilisée comme un « concept générique et lâche » qui fait référence à différentes notions telles que la transparence, la légitimité et la responsabilité (Bovens, 2007, p. 449). L'*accountability* y apparaît comme une vertu, caractérisée par les finalités qu'elle est censée poursuivre et les bienfaits qui lui sont associés. Synonyme de « bonne gouvernance », de « comportement vertueux » (Bovens, 2006, 2007, 2010; Veselý, 2013), elle s'impose de fait dans une société démocratique où la transparence s'avère de mise pour garantir la bonne marche de l'action publique et la confiance des citoyens. Les acteurs doivent rendre des comptes, et le fait d'être imputable de ses actions — *being accountable* — s'impose dès lors comme une norme qui devrait régir le fonctionnement des organisations et le comportement des acteurs (politiques, administrateurs, employés) (Bovens, 2006, 2007, 2010). Synonyme de progrès, de démocratie et de transparence, l'*accountability* devient de fait une condition *sine qua non* à la gouvernance des sociétés démocratiques contemporaines.

L'association discursive entre *accountability*, transparence (*transparency*) et responsabilité (*answerability*) permet de saisir deux aspects clefs du concept d'*accountability*. La notion de transparence place de fait la question de l'information, de sa production et de sa mise à disposition, au cœur du débat. Les citoyens doivent bénéficier de l'information nécessaire pour

des piliers de l'administration du règne. L'*accountability*, dans ce contexte, réfère à l'administration financière du royaume qui s'appuie sur le recensement des biens des sujets de la couronne et sur des mécanismes de reddition de comptes à la souveraineté qui opèrent sur une base régulière et semi-annuelle (Bovens, 2007).

apprécier et juger les actions entreprises par les gouvernements, les gouvernements doivent en contrepartie produire et garantir l'accès à cette information. De même, les États, pour être en mesure de développer une connaissance suffisante de leur administration, des organisations publiques et de leur bonne marche, doivent produire et recueillir l'information nécessaire leur permettant d'assurer la gouvernance du système.

La responsabilité entendue au sens du terme anglo-saxon d'*answerability* constitue par ailleurs un autre aspect clef de l'*accountability* dans son acception « moderne ». Elle réfère à l'idée que les acteurs doivent être en mesure de justifier leur comportement, de produire un discours qui légitime leurs actions (Ranson, 2003). La responsabilité n'est donc pas seulement entendue en termes normatifs au sens « d'obligations constitutives des individus vis-à-vis des autres membres » (Dutercq & Maroy, 2017, p. 21), mais implique que les acteurs peuvent être, et idéalement doivent être, formellement imputables de leurs actions et des effets qu'elles produisent.

1.2.3 L'*accountability* comme mécanisme

L'acception de l'*accountability* comme mécanisme (Bovens, 2010) n'est ni déconnectée de celle de l'*accountability* comme vertu, ni de sa tradition plus ancienne — *accounting* — puisque le fait de devoir rendre des comptes et d'être imputable sont centraux. En tant que mécanisme, l'*accountability* n'est néanmoins pas une vertu *per se*, mais reste associée à des externalités positives : meilleur contrôle des fonctionnements organisationnels, amélioration de l'efficacité.

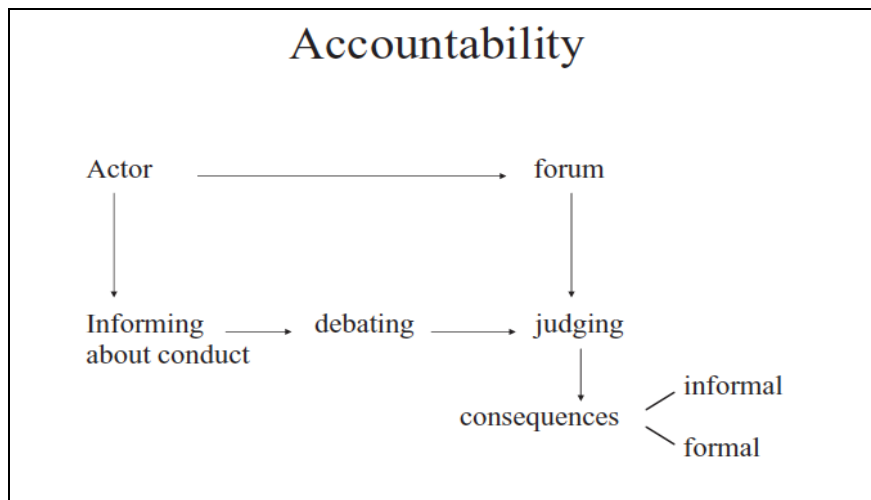
Elle réfère alors à une forme de relation sociale entre deux acteurs (individus, organisation) qui implique que le premier doit répondre de ses actes devant le second, qui bénéficie dès lors du pouvoir de juger et d'évaluer ces actions, de demander des corrections et des améliorations, et d'imposer des sanctions (Bovens, 2007, 2010; Maroy & Voisin, 2017a). Kogan (1988) en propose la définition suivante :

“Accountability can be defined as a condition in which individual role holders are liable to review and the application of sanctions if their actions fail to satisfy those with whom they are in an accountability relationship” (Kogan, 1988, p. 25)

L'*accountability* renvoie en ces termes à un mécanisme externe de contrôle de l'action qui repose sur différents éléments centraux telle que la reddition de comptes, notion très usitée dans la littérature francophone, qui n'en constitue toutefois qu'une des dimensions. Les éléments

constitutifs de ce mécanisme conceptualisé par Bovens (2010) sont illustrés dans la Figure 1.

Figure 1 : *Accountability* comme mécanisme (Bovens, 2010, p. 454)



Afin de mieux saisir les enjeux qui entourent ce mécanisme, plusieurs auteurs ont proposé d'opérationnaliser la notion au travers de questions empruntées au philosophe Wagner : *Qui doit rendre des comptes ? À qui ? Sur quoi ? Avec quelles conséquences ?* (Bovens, 2007, 2010; Leithwood & Earl, 2000). Il s'agit donc 1) d'identifier les acteurs qui font face à l'obligation plus ou moins formalisée de devoir de rendre des comptes et ceux 2) en position d'en demander ; 3) de définir les objets sur lesquels porte la reddition de comptes et 4) les conséquences associées. L'analyse basée sur ces questions opératoires permet donc de saisir les objets et les modalités de la reddition de comptes d'une part et, d'autre part, comment l'*accountability* en tant que mécanisme structure les rapports sociaux et professionnels entre les acteurs, autour d'obligations sociales ou morales, d'un rapport hiérarchique ou d'autorité formelle (Maroy & Dutercq, 2017).

Les travaux de Leithwood et Earl (2000), entre autres, qui mobilisent ces questions opératoires comme support à l'analyse des mécanismes d'*accountability* dans le champ éducatif, permettent d'éclairer le fait que l'*accountability* y recouvre plusieurs formes organisationnelles. Ces formes organisationnelles relativement traditionnelles, que nous discutons dans la sous-section suivante, tendent toutefois à évoluer vers ce que certains nomment une « nouvelle *accountability* » sous l'influence du *New Public Management* (Carnoy, Elmore, & Siskin, 2003; Carnoy & Loeb, 2002; Fuhrman, 1999; Linn, 2000; Maroy, 2013; Maroy & Voisin, 2013, 2014).

1.2.4 Formes organisationnelles traditionnelles de l'*accountability*

De longue date, en effet, les États ont mis en place des modalités de contrôle du fonctionnement des organisations scolaires basées sur des procédures de reddition de comptes. Si l'*accountability* n'est donc pas chose nouvelle dans le champ éducatif, elle présente néanmoins des formes organisationnelles différenciées qu'il est possible de saisir au travers des mécanismes de reddition de comptes, ainsi que des acteurs au cœur de ces derniers (Qui doit rendre compte? À qui?).

Dans la plupart des systèmes éducatifs, caractérisés majoritairement par une organisation de type bureaucratique (Maroy, 2008), la reddition de comptes opère traditionnellement par la chaîne administrative ou hiérarchique chargée de veiller au bon fonctionnement organisationnel selon une logique verticale *top down*. Le niveau central (État) fixe les règles et les procédures relatives au fonctionnement des instances de régulation intermédiaires (districts, commissions scolaires, académies) et des établissements scolaires, la reddition de comptes s'effectue à l'autorité administrative hiérarchique en charge du contrôle (Leithwood & Earl, 2000; Maroy & Voisin, 2013). Les relations de reddition de comptes qui lient les acteurs y sont principalement structurées par différents instruments légaux, réglementaires et budgétaires qui visent à vérifier la conformité aux règles et aux procédures fixées par l'autorité hiérarchique. Il s'agit par ailleurs de contrôler l'utilisation par les établissements scolaires des ressources allouées aux activités et aux projets dévolus (Maroy & Voisin, 2013).

Cette *accountability* hiérarchique et bureaucratique (Darling-Hammond, 2004; Leithwood & Earl, 2000; Maroy & Voisin, 2013) a pu se combiner plus récemment avec des formes organisationnelles laissant davantage de marge de manœuvre aux écoles et aux collectivités locales. Parallèlement à la mise en œuvre de politiques de décentralisation et d'autonomie scolaire, d'autres « approches » d'*accountability* (Leithwood & Earl, 2000) dites managériales ou basées sur le marché ont été introduites au sein des organisations scolaires.

Les approches managériales reposent sur l'idée que le fonctionnement des établissements scolaires devrait être « davantage orienté par des objectifs », être efficace et efficient (Leithwood & Earl, 2000, p. 14). L'accent y est donc mis sur les pratiques de gestion à la lumière d'un projet organisationnel défini. Ces approches, qui présupposent d'un lien entre mise en œuvre d'une gestion plus rationnelle et amélioration de l'efficacité des établissements scolaires, s'appuient sur

l'introduction et le développement de procédures de planification stratégique et d'orientation de l'action au sein des établissements scolaires. Des procédures administratives rationnelles planifiées, appuyées par la production de données et orientées vers des objectifs précis d'efficacité et d'efficience, devraient en ces termes contribuer à l'amélioration de l'école (Leithwood & Earl, 2000). L'expertise managériale et la valorisation de l'autonomie de gestion sont ici «le fondement du pouvoir accru de “demande de comptes” et, simultanément, d'obligation renforcée de la direction de l'école de devoir rendre compte au niveau hiérarchique supérieur » (Maroy & Voisin, 2013, p. 15).

L'*accountability* de marché (Darling-Hammond, 2004; Kogan, 1988; Leithwood & Earl, 2000) accorde quant à elle une place centrale aux mécanismes de compétition entre écoles, supposés améliorer le fonctionnement des organisations scolaires et la qualité des prestations fournies. La compétition entre écoles est encouragée par l'État qui met en place un ensemble de dispositifs — diversification de l'offre d'éducation, privatisation des établissements scolaires, publication des résultats, classement des établissements scolaires, promotion des écoles à chartes, etc. — visant à appuyer le développement d'un quasi-marché éducatif (Felouzis, Maroy, & Van Zanten, 2013; Maroy, 2006). Cette approche de l'*accountability* repose sur l'idée que les écoles sont « trop peu attentives aux besoins des élèves » (Dulude & Dupriez, 2014, p. 9), la reddition de comptes opère donc principalement des écoles à destination des parents/usagers en mesure de juger de la qualité de la prestation fournie et de choisir leur établissement. Les écoles sont néanmoins soumises à ce que Gropello (2004 cité dans Falabella, 2013, p. 25) nomme une *double accountability* puisque les autorités hiérarchiques et administratives sont aussi en position de demander des comptes aux établissements.

Au total, il s'agit de contrôler la conformité aux procédures et aux règles fixées par l'autorité hiérarchique dans une *accountability* de type bureaucratique, de porter la focale sur les modalités de gestion et de mise en œuvre du projet institutionnel dans une *accountability* de type managérial, enfin de veiller à la satisfaction des usagers dans une *accountability* de marché. Ces formes organisationnelles de l'*accountability* ne sont toutefois pas mutuellement exclusives. Avec des temporalités différentes, les systèmes d'éducation ont pu tour à tour intégrer certaines modalités organisationnelles de reddition de comptes qui peuvent, dès lors, se combiner en leur sein et évoluer au fil du temps.

Ainsi, dans les dernières décennies, des mécanismes attachés à une logique de marché et axés sur la reddition de compte au public (usagers, communauté locale) ont été introduits dans nombre de systèmes d'éducation (Ball, 2003; Falabella, 2013, 2014; Felouzis et al., 2013; Harris & Herrington, 2006). De même, des modalités de gestion managériales ont été fortement plébiscitées au sein des systèmes d'éducation et des établissements scolaires depuis les années 1980, impliquant des attentes nouvelles à l'égard des acteurs du champ éducatif en général, des chefs d'établissements et des enseignants en particulier.

Si ces formes organisationnelles de l'*accountability* sont relativement stables au fil du temps, depuis les années 2000, les objets de la reddition de comptes évoluent pour converger vers une centration sur les produits du système, plus particulièrement les résultats des élèves, des écoles et enfin des entités de régulation intermédiaires (districts, commissions scolaires, etc.). Différents instruments visant à évaluer et *monitorer* les performances du système et de ses acteurs sont progressivement introduits, parallèlement à la multiplication des modalités du contrôle organisationnel.

Cette focalisation sur les performances des systèmes d'éducation et de leurs acteurs, dès lors objets privilégiés de la reddition de comptes et du contrôle des organisations scolaires, constituent un des éléments caractéristiques du passage vers ce que certains nomment une « nouvelle *accountability* » (Carnoy et al., 2003; Carnoy & Loeb, 2002; Fuhrman, 1999; Linn, 2000; Maroy & Voisin, 2013, 2014; Webb, 2005). Cette dernière s'appuie sur un ensemble d'instruments mis en place par l'État afin de « gouverner l'éducation par les nombres » (Broadfoot, 2000; Felouzis & Hanhart, 2011; Grek, 2009; Maroy, 2013; Ozga, 2009; Rose, 1991). Elle s'inscrit dès lors dans un contexte de reconfiguration du rôle de différents acteurs et de complexification de l'action publique.

1.3 Avènement de la régulation de l'éducation par les résultats et « nouvelle *accountability* »

Les années 1990-2000 sont en effet marquées par des transformations importantes des systèmes d'éducation. Sur fond de crise économique, de recherche de compétitivité et de positionnement des pays sur le marché de l'économie du savoir (Ozga, 2009; Ozga & Grek, 2010), les systèmes d'éducation sont soumis à nombre de demandes comme celles de voir le niveau général d'éducation s'élever et « les compétences les plus demandées par l'économie se développer »

(Maroy, 2008, p. 18). L'efficacité des systèmes éducatifs est avancée comme un enjeu majeur, au cœur de la rhétorique politique. La qualité de l'éducation, considérée comme essentielle dans la construction de la prospérité économique (Ozga, 2009), est au cœur des préoccupations des systèmes contemporains. Il faut réformer l'école publique, pallier à la crise postulée des systèmes d'éducation (Boyd, 2003; Demailly, 2006).

Sous l'influence des principes de la doctrine du NPM (Hood, 1991), le rôle central de l'État dans la gestion et la production de l'éducation est par ailleurs, à partir de la fin des années 1970, peu à peu remis en question. Il s'agit de limiter les dépenses publiques, de mettre en place d'autres modes de coordination et de contrôle de l'action éducative. À la figure de l'État éducateur qui prend massivement en charge la mise en œuvre du service éducatif se substituent, avec des variations selon les contextes, celles d'un « État Régulateur »² (Majone, 1996 cité dans Maroy, 2006) devant « faire davantage place à d'autres acteurs que les autorités publiques et à d'autres formes de coordination que la Loi et la hiérarchie pour assurer les fonctions collectives » (Maroy, 2006, p. 13), et plus particulièrement de l'« État Évaluateur » (Maroy, 2006, 2008; Neave, 1988) qui définit les grandes orientations et évalue les résultats du système.

Dans ce contexte, la « nouvelle gouvernance » s'appuie sur un ensemble d'instruments qui doivent permettre de piloter le système par ses résultats. L'État central définit les programmes d'enseignement, fixe les standards et les objectifs à atteindre. Il met en place des modalités de contrôle fondées sur une évaluation *a posteriori* (Neave, 1988). Il ne s'agit donc pas pour l'État et son administration d'effectuer (uniquement) un contrôle du respect des normes et des procédures, ni de vérifier si les budgets sont bien ajustés aux besoins, mais davantage d'opérer une centration sur la mesure des produits dans une logique *ex post* (Bezes, 2005; Broadfoot, 2000; Maroy, 2006; Mons, 2009).

Il n'est donc plus seulement question de la taille des classes, du ratio maître-élèves, des ressources allouées à l'école, mais aussi des résultats et des performances des élèves, de celles des enseignants, des établissements, des districts scolaires ou du système éducatif dans son ensemble, qui font dès lors l'objet d'un contrôle accru. Par le biais de mécanismes de reddition

² Tel qu'explicité dans notre cadre théorique, la notion de régulation fait référence à des modes de régulation plus large que la régulation par la Loi au cœur de l'acceptation de la notion anglo-saxonne de « regulation ».

de comptes et d'imputabilité, on cherche à responsabiliser les agents éducatifs quant aux performances des élèves.

S'inscrivant dans une conception de la gouvernance renouvelée, cette « nouvelle *accountability* » implique que les écoles et les acteurs sont progressivement soumis à une obligation de résultats (Demailly, 2003; Lessard & Meirieu, 2004). Il faut que les apprentissages des élèves soient effectifs, mesurés par leurs performances aux évaluations (qui prennent le plus souvent la forme de tests calibrés et à grande échelle). C'est par ailleurs sur cette base que les performances des écoles et des enseignants sont comptabilisées.

Dans différents contextes, des systèmes d'*accountability* basés sur la performance (*performance-based accountability systems*) sont progressivement mis en œuvre. Ils prennent place dans les pays anglo-saxons (Australie, Nouvelle-Zélande, États-Unis, Angleterre), au Chili (Falabella, 2013), au sein des systèmes d'éducation des pays d'Europe et du reste du globe, si bien qu'il semble difficile aujourd'hui de remettre en question leur popularité auprès des décideurs politiques.

1.4 Systèmes d'*accountability* basés sur la performance, instruments privilégiés.

Les systèmes d'*accountability* basés sur la performance (SABP) peuvent être définis en nous appuyant sur les travaux de Rhoten et collègues (2003) et l'acceptation de l'*accountability* que ces derniers proposent :

“The use of tests procedure, methods, or series of tasks to measure what is taught and learned (Leithwood, Edge and Jantzi, 1999) an arrangement whereby ‘an account must be given to some authority, as an indication of compliance with defines standards, and as demonstrated by improvement on baseline or performance measure’ as determined by some form of assessment” (Rhoten et al., 2003, p. 7).

Les SABP reposent ainsi sur un ensemble d'instruments communs : fixation de standards, dispositifs d'évaluation (des acquis des élèves, des établissements scolaires) à grande échelle, mécanismes de reddition de comptes articulés à des conséquences sociales ou matérielles pour les acteurs. Portés par les politiques éducatives des dernières décennies, existants de longue date au sein des systèmes d'éducation où ils sont réactivés, alignés les uns aux autres, orientés alors par une logique instrumentale d'efficacité (Maroy, Pons, & Dupuy, 2017), ces instruments, que nous décrivons dans les sous-sections suivantes, constituent en effet le socle des SABP.

Encadré 1: Systèmes d'*accountability* basés sur la performance aux États-Unis, et la Loi fédérale *No Child Left Behind*

La politique fédérale états-unienne *No Child Left Behind* (2002) qui réautorise l'*Elementary and Secondary Education Act* constitue un cas d'école en matière de système d'*accountability* basé sur la performance et à forts enjeux (*high-stakes accountability systems* – (Carnoy & Loeb, 2002; Fuhrman, 1999; Fuhrman & Elmore, 2004). Elle s'inscrit dans la lignée d'un ensemble d'initiatives (déjà amorcées dans les années 1970 –(Frederiksen, 1994)) visant à élever le niveau global d'éducation afin de faire face aux impératifs économiques de compétitivité de la nation. Dans un contexte de revendications de certains groupes de la société civile qui dénoncent les disparités de résultats persistantes entre la majorité blanche et les minorités afro-américaines et hispano-américaines, il s'agit par ailleurs de proposer des mesures permettant de réduire l'écart de réussite entre ces groupes d'élèves.

Dans les années 1980, le rapport *A nation at Risk* (Gardner, Larsen, Baker, Campbell, & Crosby, 1983) joue tout d'abord un rôle fondamental dans l'orientation des politiques éducatives états-uniennes. Il préconise un ensemble de solutions visant à améliorer l'état du système éducatif qu'il décrit comme « médiocre » et va jouer un rôle central dans la mise en œuvre des *standards-based reforms*. Ces réformes à large échelle visent le renforcement et l'harmonisation des standards curriculaires, l'augmentation du temps d'instruction et demandent aux États des efforts pour lier responsabilités des enseignants et performances académiques (Fuhrman, 1987). Un premier pas est ainsi suggéré à l'échelle fédérale dans la mise en place de dispositifs d'imputabilité et de reddition de comptes au niveau des États :

“It was now up to them [États et gouvernements locaux] to improve education: to hold students to tougher standards, screen out unqualified teachers, reward good performance, and raise the money to do it. The federal government would stand by and cheer. It would reward excellence and teacher and schools, publish report of state and local progress, and make research findings about achieving excellence widely available” (Fuhrman, 1987, p. 136)

La promulgation de *No Child Left Behind* et la mise en œuvre de la politique du même nom, marque un tournant majeur dans l'implantation d'un système d'*accountability* à forts enjeux à l'échelle fédérale. Elle représente une intervention sans précédent du niveau fédéral dans le champ de l'éducation, relevant majoritairement de la compétence des États. La Loi instaure un ensemble de dispositifs permettant la régulation des systèmes éducatifs sur la base de leurs résultats : standards de performances, évaluation et monitoring des performances des États et des districts sur base des performances des élèves au *testing* fédéral — *National Assessment on Education Progress* (NAEP³), système d'incitatifs et de sanctions à l'égard des acteurs scolaires, imposition de dispositifs de soutien aux écoles et élèves en difficultés.

³ Fondé par l'Office pour la Recherche et l'Amélioration éducatives du Département des États-Unis de l'éducation en 1969 (année de la première administration du test, mandaté par le congrès et administré par le *National Center for Education Statistics*, l'*Institute for Education Sciences* et le Département d'Éducation [*Us Department of Education*]), le NAEP permet l'évaluation des performances des élèves des écoles publiques et privées dans plusieurs matières à différents niveaux de scolarité, à l'échelle de la nation. Si le *MAIN NAEP* permet la production d'informations sur les performances des élèves et sur les facteurs possiblement liés à ces performances, le *STATE NAEP* permet aux États de contrôler l'évolution des performances de leurs élèves au fil du temps et de faire des comparaisons entre États

NCLB instaure par ailleurs un principe de régulation des politiques compensatoires mises en place à l'échelle fédérale sur base des résultats des élèves. *Title I*, plus large programme de financement compensatoire lancé aux États-Unis en 1965 sous l'impulsion du Président Johnson dans le cadre de l'*Elementary and Secondary Education Act* (Slavin, Karweit, & Madden, 1989), prévoit l'octroi de ressources financières supplémentaires aux États et districts scolaires caractérisés par des taux importants d'élèves présentant des difficultés scolaires et issus de milieux socioéconomiques défavorisés.

Avec *No Child Left Behind* les districts et les écoles concernés par ce programme de financement compensatoire ont l'obligation de participer aux tests du NAEP. Les écoles sont contraintes d'atteindre les standards de performance fixés à l'échelle fédérale sous peine de voir les ressources allouées dans le cadre de *Title I* à terme supprimées.

Les établissements sous performants sont par ailleurs soumis à différents types de sanctions et sommés de mettre en place un ensemble de dispositifs pour améliorer leurs performances (Behrens, 2006)⁴ : soutien scolaire obligatoire pour les élèves, mise sous tutelle des établissements sous performants et intervention d'experts externes en leur sein dans une perspective d'amélioration des écoles, transferts d'élèves facilités vers d'autres établissements, fermeture des écoles avec la possibilité d'être reconverties en *Charter-Schools* (majoritairement subventionnées par des fonds privés).

L'administration Obama en demandant en 2010 la ré-autorisation de l'*Elementary and Secondary Education Act*, conserve la logique de *performance-based accountability*. Elle entend appuyer la définition de standards de performance plus ambitieux au niveau des États et s'inscrit dans la ligne directe de l'approche agressive des standards de NCLB. Le programme *Race to the Top* lancé en 2009 vise l'octroi de subventions fédérales « aux États qui proposent des standards exigeants et des politiques crédibles pour y parvenir » (Meuret, 2012, p. 78).

⁴ Behrens résume le système de régulation des performances académiques des élèves et des établissements instauré par NCLB comme suit: "States must bring all students up to the "proficient" level on state tests by the 2013-14 school year. Individual schools must meet state "adequate yearly progress" targets toward this goal (based on a formula spelled out in the law) for both their student populations as a whole as well as subgroups of students defined by socioeconomic background, race/ethnicity, and English language proficiency. If a school receiving federal Title I funding fails to meet the target two years in a row, it must be provided technical assistance and its students must be offered a choice of other public schools to attend (Title I, the largest federal education program, provides aid to low-income schools to improve education for lowachieving children). Students in schools that fail to make adequate progress three years in a row must also be offered supplemental educational services, including private tutoring and school choice. For continued failures, a school would be subject to outside corrective measures, including possible governance changes" (Behrens, 2006, p. 19).

1.4.1 Standards

La fixation de standards au niveau de l'État central joue un rôle fondamental au sein des SABP. Ces derniers permettent de définir les grandes orientations du système, constituent une norme de référence servant de repère à l'action éducative et sont largement promus dans l'idée que les « standards influent sur la rigueur académique qui, à son tour, affecte la performance »⁵ (Harris & Herrington, 2006, p. 217).

Si ces derniers peuvent porter sur des objets divers (pratiques et compétences des enseignants, niveau de qualification requis pour accéder à la profession, évaluation des établissements scolaires, etc.), c'est plus particulièrement les standards de contenus (autrement nommés standards curriculaires ou académiques) et de performance à atteindre (cibles) qui sont au cœur du débat.

Les standards curriculaires (dont l'harmonisation et la normalisation impliquent la définition d'un curriculum national) permettent de définir ce qui doit être enseigné et donc ce qui compte (Lindle, 2009). Ils impliquent une opération de hiérarchisation des contenus de savoir et offrent ainsi un guide aux activités d'apprentissage :

“Curriculum defines what counts as valid knowledge, what is knowledge, that is knowledge agreed according to the values of interested parties” (Berstein, 1975 cité dans Kogan, 1988, p. 127).

La standardisation des contenus d'enseignement a été au cœur des *Standards-Based Reforms* aux États-Unis (voir Encadré 1) dans les années 1990. Ces réformes visaient entre autres l'harmonisation des curriculums entre États ainsi que l'alignement d'un ensemble d'instruments et de dispositifs sur cette base (évaluations standardisées des acquis des élèves, contenus de formation des enseignants, etc.) :

“If governments can clearly specify what students should know and be able to do, then these standards can shape curriculum, assessment, instruction, teacher education, professional development, the allocation of resources, and all of the other elements of the educational system” (Darling-Hammond, 2006, p. 5)

En Angleterre, système traditionnellement très décentralisé, le premier curriculum national est introduit en 1988 par l'*Education Reform Act* (Broadfoot, 2000; de Landsheere, 1994). Ce

⁵ La citation originale est la suivante : *“standards affect academic rigor that, in turn, drives achievement”*

mouvement de centralisation et de normalisation des curriculums est doublé de la mise en œuvre d'un dispositif national d'évaluation visant « à mesurer les résultats obtenus [par les élèves] et ainsi à assurer la bonne mise en place de curriculum national au sein de chaque établissement » (Broadfoot, 2000, p. 48).

Les standards de performance fixés au niveau central (État ou ministère de l'Éducation) déterminent quant à eux les résultats que les élèves doivent atteindre. Par la fixation de seuils critiques en deçà desquels les résultats des élèves et des écoles sont jugés insuffisants ou inacceptables, ils introduisent la question centrale, au cœur des SABP : « *how good is good enough?* » (Lindle, 2009, p. 9). La fixation de standards de performance a ainsi pour objectif de favoriser la focalisation des écoles sur les apprentissages effectifs des élèves (Carnoy & Loeb, 2002) :

“The theory of action behind educational accountability is that setting standards and measuring performance relative to standards will lead to teachers working harder and students learning more” (Ladd & Lauen, 2010, p. 427)

Aux États-Unis, *No Child Left Behind* (Encadré 1) institue différents niveaux de performance - *proficiency levels* - au test fédéral du NAEP ainsi que les proportions acceptables d'élèves devant atteindre ces niveaux. Elle fixe comme objectifs nationaux l'atteinte pour tous les élèves des niveaux de performances exigés en lecture et en mathématiques d'ici à 2014 (Ravitch, 2010).

Les standards de contenus et de performance, couplés aux évaluations externes standardisées, occupent donc une fonction centrale au sein de la nouvelle *accountability*. Utilisés comme indicateurs de performance et de progression des élèves et/ou des établissements, ils peuvent favoriser l'analyse de certains « facteurs de réussite ou d'échec » et la proposition de « mesures de soutien [pour les élèves, les écoles] ou de développement en conséquence » (Mons & Pons, 2006, p. 19).

1.4.2 Dispositifs d'évaluation

Si l'évaluation est une pratique courante qui a accompagné la construction et l'évolution de la majorité des systèmes éducatifs, elle occupe jusque récemment une fonction essentiellement diagnostique et pédagogique (Mons, 2009; Yerly, 2014). De même, nombre de systèmes éducatifs ont mis en place des systèmes d'évaluation standardisée : c'est le cas de certains pays d'Europe comme le Royaume-Uni ou le Portugal dès les années 1950 (Eurydice, 2009), de la France dans les années 1970, ou de certains États américains.

Les évaluations externes standardisées portant sur les acquis des élèves vont néanmoins connaître une forte expansion dans les dernières décennies (Lingard, Martino, Rezai-Rashti, & Sellar, 2016). En Europe, les tests nationaux basés sur des procédures fixées au niveau central représentent, selon le rapport Eurydice (2009), une forme relativement nouvelle d'évaluation qui a pris une ampleur particulière dans les années 1990 et est aujourd'hui instituée dans une large majorité des systèmes éducatifs européens. Au Canada, ces dispositifs ont été introduits dans neuf des dix provinces relativement récemment (Ben Jaafar & Earl, 2008). Ils sont par ailleurs aujourd'hui présents dans la majorité des États américains (Linn, 2000).

Ces dispositifs d'évaluation standardisée, qu'ils soient récents ou plus anciens, voient par ailleurs leur fonction évoluer. L'évaluation permet dès lors la production de connaissances sur les systèmes d'enseignement. En testant⁶ les élèves, l'État espère obtenir une image de référence permettant de mesurer la qualité du système et d'évaluer la nécessaire mise en place d'actions pour remédier à des performances jugées insuffisantes. L'évaluation occupe donc aussi une fonction stratégique (Demailly, 2001). Elle met en lien « le politique et le pédagogique » en permettant d'une part le pilotage des systèmes éducatifs sur base de leurs résultats, d'autre part d'agir sur les mécanismes en jeu au sein des établissements, et de la classe (Mons, 2009, p. 5) :

“Testing can be used in several ways. It can be an indicator to tell administrators and teachers whether they are reaching the organization's goals and to provide information on which elements of the curriculum are reaching students and which are not. It can be used as a measure of success or failure in an incentive system. It can be used as a gauge to increase standards, to assess curricula, or to provide technical assistance. It can be used as a mechanism to allocate additional resources in order to improve outcomes for groups having difficulty reaching the standards” (Carnoy & Loeb, 2002, p. 307)

⁶ En référence à l'expression anglophone *testing*.

Finalement, les dispositifs d'évaluation (des élèves, des établissements scolaires) permettent de rendre le travail des éducateurs «de plus en plus visible au travers de l'inspection, d'observations, de la publication des taux de réussite et de performances [des élèves et des établissements] » (Webb, 2005, p. 192). Ce qui évolue parallèlement est aussi l'articulation entre dispositifs d'évaluation, mécanismes de reddition de comptes et conséquences associés à l'atteinte des objectifs fixés au niveau central (Linn, 2000). Les établissements scolaires (et les enseignants), les entités de régulation intermédiaires sont sommés de rendre compte des résultats de leurs élèves (Ben Jaafar & Earl, 2008).

1.4.3 Reddition de comptes et conséquences associées

Si nous avons déjà abordé la reddition de comptes et ses différentes formes organisationnelles, cette dernière opère dès lors au travers de divers mécanismes qui formalisent l'obligation « d'expliquer et de justifier sa conduite » (Bezes, 2005, p. 9) et les résultats qui en découlent.

À titre d'exemple, la publication des résultats des élèves, qui permet de rendre compte aux usagers des performances des écoles, est instituée dans les provinces canadiennes de l'Ontario, de la Colombie-Britannique et de la Nouvelle-Écosse (Chan, Fisher, & Rubenson, 2007), en Angleterre, dans 47 États américains (en 1999) (Fuhrman, 1999). NCLB a en outre introduit l'obligation pour l'ensemble des districts des États-Unis de publier les résultats de chaque école sous la forme de *report cards* (Behrens, 2006; Harris & Herrington, 2006; McNeil, Coppola, Radigan, & Vasquez Heilig, 2008; Reback, 2008).

De manière croissante, par ailleurs, des conséquences sociales et matérielles pour les acteurs sont articulées aux mécanismes de reddition de comptes et à l'atteinte des standards de performances. Elles exercent ainsi une pression sur les acteurs qui devraient théoriquement participer de l'amélioration de leurs performances :

“That is, if the actors in the education process are rewarded [extrinsically or intrinsically] for producing better student performance, if they are penalized for not producing high performance, they will improve performance” (Wößmann, 2007, p. 473)

Ces conséquences peuvent par ailleurs être plus ou moins formalisées, explicites (pénalités ou rétributions financières) ou implicites (symboliques), opérant dès lors moins du fait d'une action directe de la hiérarchie que par la pression de la communauté pour s'améliorer (pairs, parents, etc.) (Figlio & Loeb, 2011). Elles revêtent des formes variées, s'adressent aux acteurs à titre

individuel (les enseignants, les chefs d'établissements, les administrateurs) ou collectif⁷ (les écoles, les instances de régulation intermédiaires) et peuvent, par ailleurs, être de nature positive (récompenses) ou négative (sanctions).

Les systèmes d'incitatifs et de sanctions sont en outre l'un des éléments clefs ayant permis à la littérature de distinguer empiriquement différents types d'*accountability*. On parle d'*accountability* à « forts enjeux » ou « dure » en Angleterre et au Chili, dans certains États des États-Unis (Carnoy & Loeb, 2002; Falabella, 2013; Lee, 2008, 2010; Mintrop & Sunderman, 2009; Mons & Dupriez, 2011). Le type d'*accountability* dans ces systèmes d'éducation se distingue majoritairement par la large gamme de conséquences pour les acteurs (majoritairement négatives – (Mintrop & Sunderman, 2009), articulées aux performances des élèves : mise en probation des écoles, suppression de financements, détérioration de l'image de l'école et baisse de clientèle, etc. Dans les systèmes d'*accountability* à faibles enjeux ou douce (Maroy, 2013; Maroy & Voisin, 2014; Mons & Dupriez, 2011) les conséquences formelles associées aux performances des élèves et des écoles sont de natures différentes et reposent majoritairement sur un système de sanctions symboliques et indirectes.

Au total, si les SABP reposent sur des instruments centraux, ces derniers peuvent dans les faits être de nature et de types différents. Si nous avons choisi de décrire ces instruments de manière relativement indépendante, ils peuvent par ailleurs être combinés et plus ou moins alignés les uns aux autres (Maroy & Voisin, 2014). Ce qui est finalement en jeu dans le *design* des SABP est centralement le choix des instruments privilégiés, leur gamme, leur étendue et leur articulation.

⁷ Le rapport Eurydice (2105) distingue trois types de conséquences liées à l'évaluation des établissements scolaires en Europe : la formulation de recommandations qui doivent servir de point de départ au processus d'amélioration de la qualité ; l'obligation pour les établissements de produire un plan d'amélioration qui structure et organise les objectifs à poursuivre ; la mise en place de sanctions administrées par les évaluateurs eux-mêmes ou émanant des autorités administratives qui peuvent être prises à l'égard de l'établissement ou des responsables (diminution des ressources, retrait de certaines prérogatives, pénalités financières ou licenciement).

1.5 Stratégies de changement

Malgré la diversité empirique des outils mobilisés et les formes diverses des SABP, ces derniers reposent sur une stratégie de changement organisationnel sensiblement similaire au-delà des contextes, celle de l'alignement. Il s'agit d'améliorer les fonctionnements organisationnels par l'alignement des pratiques des acteurs en fonction d'objectifs de performance fixés au niveau central.

Les modes de régulation privilégiés au sein des SABP devraient, en effet, théoriquement réduire l'opacité du système et, plus particulièrement, le découplage entre les différents sous-systèmes composant le système d'éducation (*e.g* administration, établissement, classe), entre la « face “visible” des organisations scolaires (le programme officiel) et l'action du personnel éducatif (ce que réalisent concrètement les enseignants en classe) » (Meyer & Rowan, 1977 cités dans Barbana, Dumay, & Dupriez, 2014).

Les institutionnalistes ont en effet mis en évidence que les sous-systèmes qui composent le système éducatif (*e.g* institutionnel, administratif et technique [pédagogique]) obéissent à des impératifs différents, ce qui peut tendre vers une certaine imperméabilité des uns envers les autres. Cette idée est véhiculée au travers du concept de *loose coupling* (Weick, 1976) qui caractérise le couplage lâche entre les différentes dimensions organisationnelles (Meyer & Rowan, 1977) ou les différents paliers du système d'éducation. Par ailleurs, des écrits récents (Coburn, 2004; Rowan, 2006) ont mis en évidence le fait que ces sous-systèmes subissent toutefois l'influence des forces institutionnelles.

L'*accountability* vise donc à favoriser une meilleure cohérence organisationnelle entre les actions des différents paliers en fixant des objectifs clairs, par le biais des standards de contenus et de performances, et en prévoyant un *feed-back* devant contribuer à un effort d'ajustement. Cette stratégie de l'alignement (Lessard, 2002) s'appuie donc sur la mise en place de divers instruments permettant d'orienter les comportements des acteurs vers l'atteinte des objectifs organisationnels « dans le cadre d'une obligation de résultat pour tous » (Brassard, Lusignan, & Lessard, 2002, p. 34).

1.6 L'influence du *New Public Management*

Cette stratégie de changement organisationnel, et les instruments au cœur des SABP, ne peuvent être déconnectés de l'influence majeure de la doctrine protéiforme du NPM (Bezes, 2005) dans les tendances qui traversent l'ensemble des systèmes d'éducation. Nombreux sont les auteurs qui voient en effet dans les évolutions en cours l'intégration des principes du NPM dans la gestion et la production de l'éducation (Ball, 1997; Demailly, 2001; Lessard, 2002; Ozga, 2009; Ranson, 2003), et l'avènement d'un régime de gouvernance néolibéral (Ranson, 2003).

Le NPM articule la logique du management privé et de l'individualisation des demandes de biens publics au travers de cinq propositions principales :

« placer « le client » au sein de l'action de l'État et donc rompre avec la logique de l'administré ; décentraliser les compétences au plus près des terrains d'actions ; responsabiliser les agents de l'État sommés de rendre des comptes à la collectivité, mettre l'accent sur la qualité des services produits et l'efficacité de l'action publique et enfin, substituer au contrôle procédural traditionnel une évaluation par les résultats (Hood, 1991) » (Mons, 2007, p. 9).

Il se traduit par une « nouvelle articulation entre le social et l'économique et par une nouvelle répartition des pouvoirs et des modes d'interactions entre les gouvernements centraux et les acteurs locaux au profit de ces derniers » (Muller, 1990 cité dans Mons, 2009, p. 9).

Le NPM prône des impératifs de contrôle des bureaucraties et la rationalisation de l'action publique. Le NPM appuie par ailleurs l'idée que les « productions des services publics sont mesurables et doivent l'être grâce à des outils valides scientifiquement (...) que l'organisation publique doit ainsi être régulée par les résultats et non seulement par les contrôles procéduraux » (Mons, 2009, p. 8). Il plaide pour une « vision rationnelle et gestionnaire du contrôle des organisations » basée sur la production et le contrôle de l'information (Bezes, 2005, p. 33). Les termes du contrôle administratif sont dès lors transformés au profit de contrôles *ex post* « reposant sur le calcul et le contrôle des coûts (...) les contrôles de régularité faisant place à des contrôles d'efficacité et d'efficience » (Bezes, 2005, p. 34).

Si le NPM a alimenté les politiques de réformes de l'État depuis les années 1980 (Bezes, 2005), la stratégie néolibérale s'appuie sur la promotion des notions d'efficacité, de qualité, de transparence et de responsabilisation des agents éducatifs (Ball, 1997; Ranson, 2003). Si le référent néolibéral « a eu d'autant plus d'écho que les États étaient confrontés peu ou prou à des

contraintes et des limites effectives de financement de leurs dépenses publiques » (Maroy, 2005, p. 19), son influence se traduit par la mise en place, dans l'ensemble des systèmes d'éducation, de modes de régulation polycentriques et stratégiques qui, d'une part « ouvrent les frontières internes de l'établissement à de nouveaux partenaires conviés à participer aux processus décisionnels et à l'amélioration de la réussite scolaire », d'autre part visent à mobiliser les acteurs autour d'objectifs définis par la hiérarchie (Dulude & Dupriez, 2014, pp. 8-9).

Le développement de ces modes de régulation et de l'*accountability* dans les systèmes d'éducation s'appuie ainsi sur un ensemble de présupposés quant à leurs effets positifs sur l'efficacité et l'équité des systèmes scolaires, placées au cœur du débat public depuis les années 1990 (Hutmacher, Cochrane, & Bottani, 2001). La mise en exergue de la rhétorique qui accompagne le développement des politiques d'*accountability* aussi bien au niveau national que transnational ainsi que le rôle de certains acteurs clefs et courants de recherche dans sa diffusion nous permettront de clore cette section.

1.7 Une rhétorique commune

Enfin, au-delà des différences de contextes et malgré des modalités concrètes qui diffèrent, les politiques d'*accountability* s'appuient sur une rhétorique commune : celle de la qualité, de l'efficacité, de l'équité et de la modernité (Ball, 1997; Demailly, 2001; Lessard, 2002; Maroy, 2009; Mons, 2009; Suspitsyna, 2014). Leur développement repose sur le constat d'une double crise de légitimité et d'efficacité des systèmes éducatifs qui interagissent et se renforcent mutuellement (Boyd, 2003). Ces politiques sont ainsi envisagées comme des solutions politiques, des « remèdes à la crise des systèmes scolaires » (Demailly, 2006, p. 105), en réponse à une préoccupation grandissante autour de la qualité de l'éducation, des performances des systèmes éducatifs et de leurs résultats (Maroy, 2009). Il faut moderniser les systèmes d'éducation pour répondre aux besoins auxquels font face les sociétés contemporaines.

Les enjeux qui entourent les systèmes d'éducation, et le développement de l'*accountability* en leur sein, se situent en effet dans des préoccupations de croissance économique des pays. En lien avec la théorie du capital humain, l'éducation doit favoriser l'acquisition des connaissances et des compétences valorisées au sein des économies nationales (et de l'économie mondiale). Elle est perçue comme centrale, garante la prospérité économique. Il en va du positionnement social et sur le marché du travail des individus dans un contexte « d'importantes transformations des

marchés du travail, d'une montée du chômage, d'une précarisation plus accentuée des statuts de travail, d'une flexibilité accrue du travail, de la gestion de main-d'œuvre, une mobilité professionnelle et/ou géographique accentuée » (Maroy, 2005, p. 20).

Il s'agit par ailleurs de réaffirmer le rôle de l'école comme vecteur de cohésion sociale. Les inégalités persistantes au sein de l'institution scolaire sont vivement attaquées et l'école moderne fait l'objet de critiques acerbes que Mons résume en ces termes :

« L'institution scolaire moderne, qui devait garantir à chacun, quel que soit son statut social, une égalité des chances, n'a pas tenu ses promesses de démocratisation. En outre, la massification de l'éducation s'est traduite par le phénomène nouveau de l'échec scolaire. Au mieux, indifférente aux différences, l'école n'a pas assumé sa mission démocratique d'égalisation ; au pire, pratiquant des « discriminations négatives », qui avantagent les élèves des milieux déjà favorisés, elle participerait, en elle-même, à la production des inégalités sociales » (Mons, 2004, p. 14).

Ce discours qui place l'efficacité et l'équité des systèmes d'éducation au cœur du débat est fortement appuyé par les grands opérateurs internationaux et transnationaux (OCDE, Banque mondiale, Commission européenne) qui jouent un rôle fondamental dans la diffusion de cette rhétorique, mais aussi dans l'orientation des politiques éducatives. Au niveau international, la perspective est relativement claire, le consensus évident, dans cette affirmation extraite de la déclaration d'Incheon issue du Forum Mondial de l'Éducation⁸ :

“We are determined to establish legal and policy frameworks that promote accountability and transparency as well as participatory governance and coordinated partnerships at all levels and across sectors, and to uphold the right to participation of all stakeholders”(Forum, 2015, p. 3)

Il faut promouvoir l'*accountability* et la transparence comme mécanismes de gouvernance des systèmes d'éducation, favoriser des résultats d'apprentissages « effectifs et pertinents », développer des systèmes d'évaluation des acquis plus « précis » et responsabiliser les acteurs (Forum, 2015, p. 14).

Cet impératif fait écho à la mesure de l'éducation qui a pris une ampleur particulière dans les dernières décennies et se donne à voir, d'une part, dans les modes de régulation privilégiés au sein des systèmes d'*accountability*, et d'autre part, au travers de la popularité des grandes

⁸ Le forum mondial de l'éducation ayant donné lieu à la Déclaration d'Incheon a été tenu en 2015. Il a été orchestré par l'UNESCO et a réuni des acteurs clefs de l'éducation : décideurs politiques, représentant de la société civile, des organisations internationales et transnationales telles que l'UNICEF, l'OCDE, la Banque Mondiale (etc.).

enquêtes internationales sur les acquis des élèves (PISA – Programme International pour le suivi des Acquis des élèves, TIMS – *Trends in International Mathematics and Science Study*, PIRLS – Programme International de Recherche en Lecture Scolaire). Si ces grandes enquêtes ont vu le jour il y a une trentaine d'années, elles suscitent aujourd'hui un engouement majeur. Elles s'intègrent dans des projets d'ampleur de construction d'indicateurs internationaux permettant à la fois une comparaison entre systèmes éducatifs, mais aussi d'orienter les politiques éducatives des États-nations (Hutmacher et al., 2001).

L'enquête PISA a eu pour effet d'interpeller les pays sur la qualité de leur(s) système(s) éducatif(s), sur le degré d'efficacité et d'équité de ces derniers. Ces évaluations vont d'une part appuyer une définition de la qualité de l'éducation en termes de performances, d'autre part accroître les phénomènes de comparaison entre pays participants (72 pays membres de l'OCDE et « partenaires » pour l'édition 2015) et asseoir le rôle clef de l'évaluation standardisée dans le champ éducatif.

Chaque trois ans, la majorité des systèmes éducatifs de la planète sont évalués selon des critères d'efficacité et d'équité. Les bons joueurs sont distingués et font l'objet d'un engouement majeur, comme ce fut le cas de la Finlande, mais aussi plus récemment du Canada. Les articles de presse se multiplient, les recherches scientifiques visant à comprendre le « miracle » finlandais (Sahlberg, 2011; Simola, 2005) foisonnent. À l'inverse, les mauvais élèves sont pointés du doigt. La France est définitivement l'un des pays « développés » les plus inéquitables, « championne des inégalités scolaires » (Battaglia & Collas, 2013). Elle se situe « dans le groupe des médiocres », mais « résiste désespérément au choc PISA » (S.d., 2016). Elle ne met pas en place les réformes nécessaires.

L'*accountability* fait d'ailleurs partie des recommandations formulées par l'OCDE (OCDE, 2010, 2013b) à l'issue des tests du PISA. Elle est centrale aux réformes de la gouvernance des systèmes d'éducation appuyées par la Banque Mondiale qui mène elle aussi un projet d'envergure, s'inscrivant dans une stratégie de réformes structurelles censée amener les systèmes d'éducation à de meilleurs résultats (le projet SABER – *Systems Approach for Better Education Results*).

Il faut donc comprendre « ce qui marche », comment les systèmes d'éducation et l'école peuvent faire la différence dans la lignée des travaux de l'école efficace – *school effectiveness*. Si ce

courant a démontré empiriquement l'importance des effets classes et établissements sur les résultats des élèves⁹, le pont pour envisager l'imputabilité de l'école est alors jeté. Il postule en effet que les écoles « peuvent avoir une influence sur les performances des élèves, quelles que soient les caractéristiques sociales de leur public » (Duru-Bellat, 2002, p. 21). Il entend ainsi prendre le contrepied des conclusions du célèbre rapport Coleman (1966) et de la théorie de la reproduction sociale (Bourdieu, 1966; Bourdieu & Passeron, 1964) qui, dans le cas du premier, a mis en exergue le faible effet de l'école — *school effect* — sur les acquisitions scolaires comparativement au *background* familial des élèves et, dans le cas de la seconde, a théorisé comment l'école reproduit, voire amplifie, les inégalités sociales qui lui préexistent.

Apparu au début des années 1980, le courant de l'école efficace développe une perspective d'évaluation et de comparaison de ces derniers à une échelle nationale et internationale. Ayant participé à l'émergence du débat sur la performance des systèmes éducatifs, ce courant a par ailleurs contribué au développement de standards, de normes et de procédures permettant alors de « stabiliser les interprétations du monde scientifique et politique » (Normand, 2006, p. 37) influençant les pratiques éducatives et l'orientation des politiques d'éducation. Le projet scientifique de l'école efficace est ainsi, selon Normand (2006, p. 33), « inséparable d'un espace de production de connaissances et d'instruments de mesure visant à réduire une réalité complexe pour aider à la prise de décision politique ».

⁹Les résultats de ces travaux sont cependant soumis à controverses.

1.8 Conclusion

La première section de ce chapitre nous a permis de mettre en exergue le fait que des changements importants, transversaux à l'ensemble des systèmes d'éducation, sont en marche. Ils tracent des tendances lourdes (Lessard, 2002) qui s'inscrivent dans des transformations plus larges qui touchent l'ensemble du secteur public (Ball, 1997). Sous l'influence des préceptes du NPM, une « nouvelle régulation de l'éducation » participe de la diffusion d'un discours à caractère fortement normatif et prescriptif qui pénètre le champ éducatif, dans un mouvement général qui concerne l'État au sens large. Sous l'influence des grands opérateurs internationaux et dans un contexte où les questions de qualité de l'éducation, d'efficacité et d'équité des systèmes éducatifs sont posées avec une grande acuité, les phénomènes de comparaison entre pays s'accroissent, appuyés par les grandes enquêtes internationales, sur fond de *benchmarking* et d'imposition de standards à l'échelle internationale et nationale. La mesure de l'éducation tend alors à s'imposer comme une pratique nécessaire pour piloter les systèmes éducatifs.

L'*accountability* est ainsi envisagée comme une solution politique visant à répondre d'une part aux impératifs économiques des sociétés contemporaines, d'autre part aux préoccupations concernant les inégalités persistantes au sein de l'école. Les SABP mis en œuvre dans de nombreux systèmes d'éducation s'appuient ainsi sur une série d'instruments et servent des stratégies de changement organisationnel orientées par une logique d'efficacité, et reposent sur un ensemble de traits communs (Maroy, 2013; Maroy & Voisin, 2013, 2014) :

« [ils] “mettent en acte” et renforcent du même coup un nouveau paradigme politique où l'école est conçue non plus comme une institution, mais comme un “système de production” (Maroy & Mangez, 2011) ; les objectifs opérationnels sont exprimables par des données quantifiées, qui deviennent des “standards” auxquels on confronte les résultats effectifs du système — on gouverne ainsi par des “nombres” (Rose, 1991 ; Ozga, 2009) ; les instruments divers d'évaluation des acquis des élèves (testing) sont centraux, même si les modalités de mise en œuvre et d'usage de ces outils d'évaluation peuvent être très variables en fonction des contextes ; divers outils d'action publique (contractuels, financiers, réglementaires) organisent les “conséquences” de la reddition de comptes pour les acteurs individuels ou collectifs aux différents paliers du système » (Maroy & Voisin, 2014, p. 35)

Si les SABP reposent sur des théories du changement ((Maroy & Voisin, 2014) qui, comme nous le verrons plus avant, peuvent être sensiblement différentes en fonction des formes qu'ils revêtent et des dispositifs qui y sont priorisés, ils reposent toutefois sur un ensemble de présupposés communs. Ces présupposés ont trait aux potentialités des SABP et de leurs instruments centraux que l'on peut résumer de la sorte : la définition de standards permet une définition claire des objectifs à atteindre pour les enseignants et les élèves, l'évaluation standardisée produit quant à elle une image de référence quant aux apprentissages effectifs des élèves et, par extension, de l'efficacité des pratiques d'enseignement mises en œuvre. Elle éclaire ainsi la nécessité de mettre en place des mesures de remédiation le cas échéant (pour soutenir les apprentissages des élèves ou le développement professionnel des enseignants). La reddition de comptes et les incitatifs attachés aux résultats permettent de responsabiliser les acteurs. Ils induisent par ailleurs une pression sur ces derniers pour qu'ils améliorent leurs performances et mobilisent les ressources à disposition pour ce faire (Mons, 2009). Ces mécanismes fonctionneraient ainsi « de concert dans une dynamique positive, [qui] permettraient de rendre le système éducatif plus efficace » (Mons, 2009, p. 108).

Par là même, ces dispositifs auraient un effet positif sur l'équité scolaire :

« En imposant des standards communs à tous les élèves, ces mesures obligent les enseignants à développer des attentes uniformes à leur égard, quelles que soient leurs caractéristiques individuelles (handicap, appartenance à une communauté ethnique minoritaire ou à un milieu social défavorisé) et quels que soient les contextes scolaires dans lesquels ils évoluent (écoles accueillant ou non un public favorisé). Ceci aurait un effet pédagogique « homogénéisateur » (même exposition aux contenus d'enseignement, même nombre d'heures utiles, attentes face aux élèves identiques), qui tendrait à limiter les inégalités scolaires globales et celles d'origine sociale, en permettant d'améliorer les résultats des élèves appartenant aux groupes sociaux considérés comme défavorisés » (Mons, 2009, p. 109)

Une littérature, somme toute récente, mais prolifique, permet de confronter ces présupposés à l'épreuve de l'empirie. Depuis les années 2000, un nombre conséquent de recherches économétriques visant à mesurer les effets des SBAP, d'une part sur l'augmentation des performances des élèves et des écoles (efficacité scolaire), d'autre part sur les écarts de réussite entre groupes d'élèves (équité), s'est en effet développé, entre autres dans le contexte états-unien. Ces études font suite à la mise en œuvre des *Standard-based reforms* (SBR), de NCLB à l'échelle fédérale (Encadré 1) et des SABP au sein des systèmes éducatifs des États-Unis.

Ces travaux permettent d'interroger et de confronter la rhétorique qui entoure leur développement à l'aune de résultats factuels : les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ? C'est aux résultats et aux débats soulevés par ces études que nous nous intéressons dans la seconde section de ce chapitre de problématique.

SECTION 2 : Effets des systèmes d'*accountability* basés sur la performance sur l'efficacité et l'équité scolaires, revue de la littérature

Cette section propose une synthèse des travaux qui visent à évaluer les SABP à la lumière de critères d'efficacité et d'équité scolaires. Les recherches qui y sont discutées, menées majoritairement dans le contexte de Etats-Unis et par des économistes, adoptent tantôt une perspective comparative inter-États, tantôt une perspective « nationale ».

Une stratégie de sélection, pour proposer une synthèse de la littérature qui s'intéresse aux effets des SABP sur l'efficacité et l'équité scolaires, consiste à porter la focale sur les études comparatives corrélationnelles ou causales inter-systèmes, utilisant des données qui permettent de comparer leurs résultats. Lee (2008), dans une rigoureuse méta-analyse, identifie une dizaine d'études états-uniennes sur base de ces critères d'inclusion. Mobiliser ces dernières comme point de départ de notre synthèse apparaît d'une grande pertinence pour notre propos. En effet, la discussion de leurs conclusions nous permet de mettre en exergue le fait que les tendances qui émergent de ces études, qui font office de références en la matière, font toutefois l'objet de débats, de controverses, tantôt remises en question, tantôt appuyées par les études subséquentes.

Ces études constituent donc le cœur de notre synthèse, même si nous faisons par ailleurs référence aux résultats de certaines études comparatives internationales, à des revues de littérature et de méta-analyses ainsi qu'à des études nationales. Nous discutons donc tout d'abord dans cette sous-section les conclusions des études analysant les effets des SABP sur l'efficacité scolaire. Nous nous penchons par la suite sur la question de l'équité, pour aborder celle des effets imprévus ou inattendus des SABP. Nous présentons finalement les perspectives générales de notre recherche.

1.9 De l'efficacité

“International evidence shows a strong positive association between the quality of schooling, measured by students’ performance on cognitive achievement tests, and long-run economic growth across countries” (Barro, 2001 ; Hanushek, 2005 ; Hanushek & Kimko, 2000 ; Wößmann, 2003 cités dans Wößmann, 2005, p. 473)

Si l'éducation est reconnue comme l'un des principaux vecteurs de compétitivité et de croissance économique, l'efficacité des systèmes d'éducation apparaît, nous l'avons vu, comme un enjeu majeur au cœur de la mise en œuvre des SABP. Les études mobilisées dans cette sous-section s'intéressent aux effets de certains instruments centraux des SABP sur les résultats des élèves (*testing* portant sur l'acquisition des compétences minimums, examens de promotion et de graduation, enjeux attachés aux résultats des élèves).

1.9.1 Effets de l'introduction du *testing* sur l'efficacité scolaire, les études comparatives inter-États (États-Unis)

L'étude de Fredericksen (Fredericksen, 1994) pour l'*Educational Testing Service* de l'*US Department of Education* est pionnière en la matière (Lee, 2008; Mons, 2009). Elle s'intéresse aux effets de l'introduction des SABP, plus particulièrement du *testing* portant sur les compétences minimum¹⁰s (*Minimum Competencies Tests – MCT*), sur les apprentissages des élèves et sur les pratiques enseignantes¹¹.

Fredericksen compare les performances des élèves des écoles publiques de trois groupes d'âge (9 ans, 13 ans et 17 ans) au NAEP (voir notre de bas de page 3), dans 48 États, sur deux temporalités : 1976 et 1984. Il adopte une logique de comparaison pré et post introduction des MCT au sein des États, qu'il différencie en fonction des enjeux associés aux performances des élèves. Cette catégorisation des États à partir des enjeux liés aux MCT est construite sur base des travaux de Goertz (1986; 1988 cités dans Fredericksen, 1994, p. 14) et réfère principalement au caractère obligatoire des MCT et aux enjeux pour les élèves (suivi des résultats et mise en place

¹⁰ Le *Texas Assessment of Basics Skills*, exemple de MCT, est introduit au sein de l'État en 1980. Il deviendra, dans les années 1990, le pivot du SABP à forts enjeux texan, ayant fortement inspiré celui qui a été institué par NCLB au niveau fédéral.

¹¹ Ces tests (MCT) sont introduits aux États-Unis à partir des années 1980 et sont au cœur des SBR, dans la lignée du « *back to basics mouvement* » initié dans les années 1970 (Fredericksen, 1994, p. 9). Dans une logique de focalisation sur les fondamentaux (numératie, littéracie), ils déterminent un seuil de connaissances et de compétences minimums que tous les élèves doivent posséder à l'issue de niveaux de scolarité donnés pour devenir des « citoyens productifs » (Amrein et Berliner, 2002, p. 3).

d'action de remédiation, promotion vers le cycle supérieur ou certification assujettie à l'atteinte des standards de performance aux MCT) (Frederiksen, 1994).

L'étude analyse les performances des élèves en comparant les résultats de ces derniers aux questions visant à évaluer d'une part des compétences de base (*routine items*), d'autre part des compétences plus complexes (*non routine items*)¹². Friedericksen conclut à une augmentation significative et plus importante des résultats des élèves dans les États à « forts enjeux », les gains étant en outre plus importants dans les items évaluant les compétences de base ainsi que pour les groupes d'élèves les plus jeunes (9 ans). L'étude suggère par ailleurs que les MCT auraient davantage d'effets lorsqu'assortis d'enjeux pour les élèves. L'auteur arrive aux conclusions que, pour les élèves les plus âgés (13 ans et 17 ans), la focalisation sur l'enseignement des compétences de base nuirait par ailleurs à l'acquisition de connaissances plus complexes :

“It seems reasonable to conclude that the use of MCTs can have desirable influences on the performance of young students as measured by NAEP, especially when high-stakes conditions prevail. It also seems reasonable to assume that for teenage students too much emphasis on teaching basic skills may indeed interfere with the teaching of higher-order thinking skills. If teaching basic skills interferes with acquiring nonroutine skills, they would surely interfere with teaching more advanced thinking abilities” (Fredericksen, 1994, p. 20)

Bishop et collègues (2001) s'intéressent aux MCT ainsi qu'aux examens de fin de session (notre traduction littérale de *End of Course Exams – ECOCE*) qui sanctionnent l'acquisition des connaissances et des compétences dans certaines matières dispensées, à l'issue d'une période de formation donnée¹³. Cette étude s'appuie d'une part sur les données du NAEP, d'autre part sur celles du NELS (*National Education Longitudinal Study*¹⁴), et intègre à l'analyse des effets des

¹² Fredericksen définit ces deux types de compétences comme suit : les compétences de base réfèrent aux “*elementary abilities necessary for the lower levels of performance in such areas as reading, writing, and arithmetic*” et les compétences de haut niveau renvoient à celles qui sont nécessaires “*for a deeper understanding of literature, science, math, etc.*” (Frederiksen, 1994, p. 12).

Les hypothèses au cœur de la recherche sont formulées de la sorte : “*One hypothesis was that the use of MCTs would result in improvements in the learning of basic skills - elementary abilities necessary for the lower levels of performance in such areas as reading, writing, and arithmetic. A second hypothesis was that any improvement in basic skills would be gained at the expense of higher-order skills - those needed for a deeper understanding of literature, science, math, etc.*” (Frederiksen, 1994, p. 12)

¹³ Les *California's Golden States Exams*, introduits dans les années 1990, ou les bien plus anciens *Regent Courses and Exams* de l'État de New-York en sont des exemples. Ils sont fortement axés sur les curriculums nationaux et présentent des niveaux d'exigences plus élevés que les MCTs.

¹⁴ Le NELS est une étude longitudinale du *National Center for Education Statistics* portant sur les acquisitions des élèves scolarisés en 8^e année aux États-Unis, âgés de 13 à 14 ans. L'édition 88 (NELS : 88) qui porte sur un échantillon national représentatif a été suivie par quatre éditions en 1990, 1992, 1994 et 2000. Au-delà de tests

MCT les incitatifs qui leur sont associés, dans une logique quelque peu différente de celle de Fredericksen. Les enjeux associés au *testing* pour les établissements scolaires et les enseignants (« *SBR strategy* » : publication des résultats, récompenses, sanctions) sont en effet aussi pris en compte.

Bishop et collègues aboutissent aux conclusions que les MCT n'auraient en soi qu'un faible impact sur les performances des élèves mesurées au *testing* fédéral, excepté lorsqu'ils sont couplés avec des ECOCE. Ces effets seraient en outre plus importants pour les niveaux de scolarité les plus bas et lorsque ces dispositifs sont associés à des incitants. Ils arrivent ainsi à des conclusions relativement similaires:

MCE also have modest effects on average levels of achievement. They do, however clearly have desirable distributional consequences for they focus their incentive effects on the school's struggling or lagging students. When they are combined with stakes for teachers and schools, as they typically are, the package has very respectable effects on test scores. (Bishop, 2001, p. 61)

Jacob (2001) se focalise quant à lui sur les examens de graduation du secondaire supérieur, autrement nommés examens de sortie du secondaire (*High-School Examination Exit Exams – HSEEE*), qui visent à vérifier l'acquisition des fondamentaux à la sortie de ce niveau d'études et aboutissent généralement à la délivrance du diplôme de fin de secondaire.

L'étude de Jacob aboutit à des conclusions sensiblement différentes quant aux effets des examens de sortie du secondaire sur les performances des élèves. Mobilisant les données du NELS, la comparaison inter-États intègre par ailleurs les variables de genre, d'origine socioethnique (blancs et afro-américains) et de milieu socioéconomique¹⁵ comme contrôle. Le différentiel de gains de performances pour les élèves situés dans les derniers déciles et quartiles de la distribution et les gains en termes de performances moyennes sont par ailleurs questionnés.

Selon Jacob, les HSEEE n'auraient pas d'impact sur la moyenne des scores des élèves, une fois les acquis antérieurs des élèves pris en compte. Il conclut finalement à une association forte entre ces examens et le décrochage scolaire. Selon cette étude en effet, 25 % des élèves situés dans les

administrés aux élèves (en mathématiques, sciences, lecture, humanités), l'enquête prévoit une série de questionnaires pour les élèves, les enseignants, les administrateurs scolaires et les parents. Pour plus d'informations <https://nces.ed.gov/surveys/nels88/>.

¹⁵ La variable composite SES incluse aux données du NELS incorpore le revenu de la famille, la profession des parents et leur niveau d'éducation (Jacob, 2001, p. 106)

déciles les plus bas de la distribution seraient plus à même de décrocher dans les États ayant mis en place des HSEEE (Jacob, 2001).

1.9.2 Effets des SABP à forts et faibles enjeux, les études comparatives inter-États

En ce qui concerne l'effet des tests à forts enjeux, plus particulièrement l'effet des examens de sortie du secondaire (HSEEEs), l'étude longitudinale à partir des données du NAEP¹⁶ d'Amrein et Berliner (2002) aboutit à des conclusions proches de celles de Jacob. Si ces auteurs différencient les systèmes d'éducation en fonction des enjeux associés aux tests, ils construisent un indice de pression prenant en compte un ensemble d'instruments associés aux résultats des élèves à ces examens et leurs résultats (publication des résultats, sanctions pour les écoles, etc.), différenciant ainsi les systèmes d'*accountability* à forts et à faibles enjeux.

Les auteurs arrivent ici aux conclusions d'effets inconsistants des SABP à forts enjeux sur les performances moyennes des élèves. Si des gains sont constatés dans les États prévoyant des HSEE couplés à des enjeux forts pour les établissements scolaires, ils sont peu stables (tantôt positifs tantôt négatifs) et seraient liés à des phénomènes d'exclusion des élèves des cohortes soumises au NAEP. De fait, les auteurs, dont les conclusions rejoignent celles de Jacob (2001), émettent de sérieuses réserves sur les politiques étudiées :

“At the present time, there is no compelling evidence from a set of states with high-stakes testing policies that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators. Because of this, the high-stakes tests being used today do not, as a general rule, appear valid as indicators of genuine learning, of the types of learning that approach the American ideal of what an educated person knows and can do” (Amrein et Berliner, 2002, p. 57).

Carnoy et Loeb (2002) conservent cette logique de comparaison des performances en fonction du degré de pression exercé par les États sur les établissements scolaires en lien avec le *testing*. Leur indice de pression, qui fait office de référence en la matière, est construit à partir de l'analyse des données du *Consortium for the Policy Research in Education's survey* (données autorapportées

¹⁶ Mais aussi les données du SAT et ACT: *“The SAT Reasoning Test is a standardized test that assesses the critical reading, mathematics, and writing skills that students need to be successful in college. Each of the three sections that comprise the SAT Reasoning Test has a possible score of 800 points. Prior to 2005, the SAT test included only two sections, the verbal section (now referred to as the critical reading section) and the math section, each having possible scores of 800 points. SAT test results represent one factor considered by many colleges and universities in making admissions decisions. The SAT is owned, published, and developed by the College Board, a non-profit organization in the United States”* (<http://www.cde.ca.gov/ds/sp/ai/whatissat.asp>).

recueillies par questionnaires). Il prend en compte les instruments et les incitatifs mis en place au niveau des États pour réguler le fonctionnement des écoles, suivre et contrôler les performances des élèves (fixation de standards de performance, fréquence du *testing* et niveaux de scolarité testés, sanctions et récompenses pour les écoles, etc.).

L'ensemble des États est ainsi placé sur un continuum (échelle à cinq points) qui distingue les SABP à faibles et à forts enjeux —*high stakes vs low stakes*. Ils arrivent aux conclusions que les États les plus peuplés, avec une forte proportion d'élèves issus de minorités socioéconomiques et ethniques, sont plus à même de mettre en œuvre des SABP à forts enjeux. Ils concluent par ailleurs à la présence d'une relation forte, statistiquement significative entre la force des enjeux et les gains de performance en mathématiques, pour les niveaux de scolarité les plus élevés. Leurs résultats sont donc différents de ceux d'Amrein et Berliner, mais ne prennent toutefois pas en compte la question du décrochage scolaire, ce qui pourrait expliquer ces divergences.

Hanushek et Raymond (2005) qui conservent la logique de distinction des SABP en fonction des enjeux associés, analysent l'impact de divers instruments (publication des résultats des élèves, incitatifs pour les écoles et les enseignants, etc.), mais ne mobilisent pas d'indicateur composite pour opérationnaliser l'*accountability*. L'étude longitudinale multiniveau, menée à partir des données du NAEP, aboutit aux conclusions que les SABP à forts enjeux sont effectivement liés à une augmentation des performances des élèves. Néanmoins, différents instruments visant à pressuriser les écoles doivent être combinés. La publication des résultats sous la forme de *report cards* est à elle seule insuffisante et doit être assortie d'autres conséquences pour les établissements scolaires pour que les gains de performance puissent être observés :

“We find that the introduction of accountability systems into a state tends to lead to larger achievement growth than would have occurred without accountability. The analysis, however, indicates that just reporting results has minimal impact on student performance and that the force of accountability comes from attaching consequences such as monetary awards or takeover threats to school performance. This finding supports the contested provisions of NCLB that impose sanctions on failing schools” (Hanushek et Raymond, 2005, p. 298)

Ces résultats sont dans une certaine mesure confirmés, bien que fortement nuancés, par Lee et Wong (2004) qui différencient les systèmes éducatifs états-unien selon deux catégories — *accountability* dure (*strong*) et *accountability* faible (*weak*). Mobilisant les données du NAEP (1990, 1992, 1996, 2000), ils mettent eux aussi en exergue une augmentation de la moyenne des

scores des élèves en maths pour les niveaux les plus élevés (8^e). Toutefois, les gains de réussite positivement liés à « l'activisme » des États dans l'*accountability* (*accountability* dure) ne se maintiendraient pas à long terme.

1.9.3 Le cas texan

Si la question des enjeux pour les acteurs et de leur impact sur les performances des élèves est centrale aux études comparatives inter-États discutées jusqu'ici, plusieurs études « nationales » ont été menées sur cette question (voir (Chiang, 2009) pour la Floride (Ladd & Lauen, 2010), pour la Caroline du Nord). Plusieurs d'entre elles ont par ailleurs été réalisées dans le contexte du Texas afin de comprendre les gains de performance fulgurants réalisés par cet état à la suite de l'introduction d'un SABP à forts enjeux. L'étude de Reback (2008) permet de mettre en lumière certaines stratégies mises en place au sein des établissements scolaires et susceptibles d'expliquer l'amélioration des performances constatées.

S'appuyant sur les données du *Texas Assessment of Academic Skills* (TAAS), il aboutit effectivement aux conclusions d'une augmentation des résultats des élèves. Si ces gains s'observent à court terme, l'étude met en exergue, au travers d'une stratégie de *tracking* des performances des élèves mises en lien avec les systèmes d'évaluation et de notation des écoles, le fait que ce sont les sous-groupes d'élèves dont la performance s'avère cruciale dans l'évaluation des performances des écoles qui réalisent des gains notables. Du fait de la pression attenante au système d'évaluation des performances des écoles (*ranking*), les établissements scolaires et les enseignants opéreraient un changement drastique dans l'affectation des ressources éducatives entre élèves et selon les contenus d'enseignements dits à enjeux. Autrement dit, afin d'atteindre les standards de performance, les écoles émettraient des directives quant aux matières et aux contenus prioritaires et les enseignants se concentreraient quant à eux davantage sur les élèves « à la marge » des seuils de performance, au détriment des élèves les plus performants :

“A greater incentive to improve the performance of a particular student in math decreases that student's performance in reading, and a greater incentive to improve the reading score decreases a student's performance in math. Within a school in a given year, distributional effects are amplified for grades in which there are relatively strong incentives to improve performance. For example, for math achievement in the relatively important grades, low achieving students perform better and relatively high achieving students perform worse than similar schoolmates in other grades. There is also a positive relationship between short-run, school-level accountability incentives and the performance

of students in the terminal grades of their schools” (Reback, 2008, p. 1412)

On peut ainsi déduire des conclusions de Reback que les gains de performance enregistrés ne seraient finalement que le reflet de pratiques d’entraînement intensif des élèves aux tests, hypothèse par ailleurs formulée par plusieurs auteurs (Linn, 2000; Mons, 2009). Qui plus est, les gains de performance observés dans les tests à l’échelle des États ne se retrouvent généralement pas aux tests intervenant à l’échelle fédérale (Lee, 2007; Linn, 2000). L’ensemble de ces éléments vient ainsi remettre en question le miracle texan.

1.9.4 Effets des SABP, les études comparatives internationales

Qu’en est-il des résultats des études comparatives internationales qui portent sur l’*accountability* ?

Wößmann (Wößmann, 2005, 2007) fait en quelque sorte office de référent du fait du peu d’études menées. Il utilise les données des grandes enquêtes internationales (TIMSS 1995 et TIMSS Repeat 1999; PISA 2000) pour questionner l’effet de la compétition entre établissements publics et privés, de l’autonomie scolaire et de l’*accountability* sur les performances des élèves. L’*accountability* dans cette étude fait plus particulièrement référence à la présence d’examens externes standardisés, dans la même veine que le MCT ou les ECOCE discutés plus haut.

S’il aboutit aux conclusions d’un effet positif de la présence d’examens ou de tests centralisés, Mons (2007) nuance ces résultats. À partir des données du PISA 2000, elle en vient à la conclusion que les effets constatés dans l’étude de Wößmann auraient davantage trait au niveau de développement économique des pays. Une fois que cette variable est intégrée (comme variable de contrôle), les effets d’abord mis en exergue par les analyses s’annulent (Mons, 2009).

L’OCDE (2007) arrive aussi à des résultats mitigés en ce qui a trait aux effets de certains dispositifs sur les performances des élèves (information aux parents sur les performances des élèves, usage des données pour évaluer les enseignants). D’après les résultats de leurs analyses multiniveaux, la publication des résultats des élèves est le seul instrument qui demeure significatif après que certains facteurs socioéconomiques et démographiques aient été pris en compte.

En résumé, qu’en est-il des effets des SABP sur l’augmentation des performances des élèves ? S’il semble difficile d’apporter une réponse univoque à cette question, la tendance qui semble

néanmoins émerger des études états-uniennes est celle de la participation des SABP à forts enjeux à des performances plus élevées dans certaines matières (les mathématiques) et pour les niveaux de scolarité les plus élevés (8^e plus particulièrement).

La méta-analyse de Lee (2008) apporte toutefois des nuances à ces tendances qui apparaissent déjà relativement fragiles. Ce dernier, après avoir analysé un ensemble de paramètres centraux des principales études causales et corrélationnelles comparatives inter-États américains (opérationnalisation de l'*accountability*, variables entrées à l'analyse, échantillon, unités d'analyses, données, méthodes d'analyse des données, analyse des tailles d'effets¹⁷), discutées pour la plupart dans cette sous-section, arrive au constat d'évidences relativement faibles quant au lien entre *accountability* et augmentation des performances des élèves. Une tendance s'observe en effet en ce qui a trait à des gains de performance (Figlio & Loeb, 2011; Lee, 2008; Mons, 2009). Toutefois, ces effets demeurent relativement faibles (voir ANNEXE 1) et peu stables dans le temps. Qui plus est, des résultats divergents et des considérations méthodologiques imposent d'user de précaution avant de pouvoir envisager une quelconque généralisation des résultats de ces études (Linn, 2000; Lee, 2008).

1.10 De l'équité

An Act - To close the achievement gap with accountability, flexibility, and choice, so that no child is left behind (Congress, 2002)

Aux États-Unis, l'écart persistant de réussite entre les élèves issus de la majorité blanche et ceux des minorités afro-américaines et « hispaniques » est jugé inacceptable. L'objectif de réduction de l'écart de résultats entre groupes d'élèves et d'amélioration de la réussite académique des groupes les plus défavorisés ont été au cœur de la mise en œuvre de NCLB (Carnoy & Loeb, 2002; Harris & Herrington, 2006). De fait, c'est plus particulièrement la participation des SABP à la diminution de l'écart de réussite entre groupes d'élèves, défini selon des critères socioéthniques, qui a fait l'objet de recherche aux États-Unis.

Les études de Carnoy et Loeb (2002) et Hanushek et Raymond (2005), déjà discutées dans la sous-section précédente, s'intéressent aux effets des SABP à forts enjeux sur l'écart de

¹⁷ Nous avons placé en annexe (ANNEXE 1) un tableau de synthèse qui précise les tailles d'effets des études discutées dans ces sous-sections qui reprend en partie des éléments des travaux de Lee (2008).

performance entre les élèves issus de la majorité blanche et les élèves afro-américains et hispaniques. Ils aboutissent sur ce point à des conclusions qui, en partie, divergent.

Carnoy et Loeb (2002) concluent en effet à des gains de performance pour tous les groupes d'élèves, au moins dans l'atteinte des niveaux de performance de base en mathématiques au NAEP. Ces gains s'avèrent en outre plus importants pour les élèves hispaniques et afro-américains, plus particulièrement dans les niveaux de scolarité supérieurs et dans les SABP à forts enjeux.

“Gains for other racial/ethnic groups from greater emphasis on student outcomes and accountability are even greater, perhaps due to lower rates of basic proficiency in 1996”
(Carnoy & Loeb, 2002, p. 313)

Puisque plusieurs études ont par ailleurs associé les gains de performance à l'exclusion des élèves les plus fragiles du *testing* (dont celles d'Amrein & Berliner et de Jacob discutées plus haut), Carnoy et Loeb (2002) intègrent donc les taux de redoublement et d'exclusion comme contrôle dans leurs analyses. S'ils ne peuvent conclure à un impact des SABP sur les taux de redoublement et la progression dans les niveaux supérieurs (après la 8^e année) des élèves blancs américains et afro-américains, ils précisent néanmoins ne pouvoir exclure la possibilité d'une augmentation des taux de redoublement pour les élèves hispaniques (Carnoy & Loeb, 2002).

Si Hanushek et Raymond (2005) appuient le constat de gains pour tous les élèves, il apparaît toutefois que les améliorations des résultats de ces deux groupes (les élèves afro-américains et hispaniques) sont nettement inférieures à ceux de la majorité blanche. Plus particulièrement, l'écart de performance entre élèves afro et blancs américains se creuse davantage dans les SABP à forts enjeux. Dans un contexte d'augmentation de la ségrégation scolaire, les effets de composition, qui renvoient à la concentration d'élèves issus des minorités ethniques et socioéconomiques au sein des établissements scolaires, affectent davantage les élèves afro-américains qui cumulent plusieurs handicaps. Les auteurs sont donc plus critiques que les précédents dans leurs conclusions :

“The introduction of consequential accountability systems has a clearly beneficial impact on overall performance. But other forces are simultaneously pushing the distribution of performance— particularly as observed in the Black-White achievement gap—in less desirable ways. First, accountability as seen during the 1990s tended to help White achievement more than Black achievement. Second, the observed movement toward higher minority concentrations in schools has a detrimental effect on Black achievement, again pushing toward a wider distribution of achievement” (Hanushek & Raymond, 2005, p. 321)

Lee et Wong (2004) arrivent cependant à des conclusions plus nuancées. Pour ces derniers, aucun effet différencié des SABP à enjeux forts versus les SABP à enjeux faibles n’est constaté sur la réduction de l’écart de réussite entre ces trois groupes d’élèves. Il en va de même lorsque la différenciation des groupes est opérée non pas en termes ethniques, mais selon des caractéristiques socioéconomiques. Ces résultats témoignent donc, selon les auteurs, de l’incapacité des SABP à tenir leurs promesses.

Au niveau international, si Schütz, West et Wößmann (2007) concluent à un effet positif des examens externes standardisés sur les performances de tous les élèves, ils reconnaissent toutefois que les gains sont moins importants pour les élèves les plus fragiles, les élèves ayant un niveau socioéconomique plus élevé tirant davantage profit de ces dispositifs. Ces auteurs s’intéressent par ailleurs aux effets d’interaction entre certains dispositifs d’*accountability* mobilisés au sein des établissements scolaires et le niveau socioéconomique des élèves : usage de tests calibrés et des résultats des élèves pour prendre des décisions quant à leur promotion, pour suivre les progrès des écoles au fil du temps ou comparer leurs performances à celles des autres écoles. Leurs analyses aboutissent ici à des résultats peu concluants. Ces dispositifs n’engagent aucun effet différencié en fonction du niveau socioéconomique des élèves.

Quelles conclusions tirer de la synthèse de ces études ? Il apparaît que, malgré des résultats mitigés, les SABP auraient tendanciellement un effet négatif sur le différentiel de performances entre élèves, plus particulièrement lorsque ces derniers sont distingués sur base de leur origine ethnique, dans le contexte des États-Unis.

Au total donc, si les auteurs cités ici tendent à être plus optimistes bien que nuancés en ce qui a trait aux effets potentiellement positifs des SABP sur l’efficacité scolaire, ils sont davantage sceptiques quant à leur capacité à réduire l’écart de performance entre groupes d’élèves. Les conclusions et les recommandations qu’ils formulent vont par ailleurs dans des sens relativement

différents.

Pour Hanushek et Raymond (2005), le défi de la réduction de l'écart de performance entre groupes d'élèves reste entier. Néanmoins, puisque les SABP produisent des gains de performance pour certains groupes d'élèves, il faut que ces derniers convergent autour de la mise en œuvre d'incitatifs mieux ciblés pour les écoles.

Amrein et Berliner (2002) ainsi que Lee (2008, 2010) sont plus critiques et remettent frontalement en cause l'approche agressive par les standards au cœur de la politique fédérale NCLB et des SABP à forts enjeux. Reposant sur la fixation de standards irréalistes et sur une perspective d'amélioration à court terme, l'approche prônée au niveau fédéral limiterait la capacité des acteurs à s'engager dans un processus de réflexion et dans la mise en place de solutions viables à long terme, capables de produire une réelle amélioration de la qualité de l'enseignement et des apprentissages des élèves (Lee, 2010).

La pression induite par des dispositifs de contrôle couplés à des standards trop contraignants et l'accent mis sur les sanctions davantage que sur les mesures de soutien (Amrein & Berliner, 2002; Mintrop & Sunderman, 2009) pousseraient par ailleurs les acteurs à mettre en place des stratégies visant à déjouer le système. Bien qu'elles ne constituent pas le cœur de notre propos, il nous semble pertinent de faire brièvement référence à ces stratégies et donc aux effets « imprévus » des SABP qui permettent d'observer dans quelle mesure les acteurs mobilisent leur marge de manœuvre, au demeurant fortement réduite par l'introduction des SABP, pour répondre aux pressions et aux exigences institutionnelles.

1.11 Effets imprévus

Sur ce point, une autre étude de Lee (2010) est particulièrement intéressante. Ce dernier modélise et analyse les stratégies mises en place par les États pour répondre à la pression imposée par le niveau fédéral par le biais de NCLB. Puisque les États bénéficient d'une certaine marge de manœuvre dans la mise en œuvre des SABP, il fait l'hypothèse que ces derniers répondent aux demandes et pressions institutionnelles par la mise en œuvre de stratégies d'adaptation et de « survivance », dont certaines ont vocation à déjouer le système. Ces stratégies de *gaming* (Hood, 2007) permettraient donc de contourner les impératifs, tout en maintenant l'illusion d'une certaine cohérence institutionnelle :

“*Observed survival strategies include goal reduction (e.g., lowering cut scores for proficiency standard), the procrastination of policy implementation (e.g., delaying the schedule of proficiency target raises for schools), and exclusion of low-achieving students from testing or reporting, all of which threaten the validity of test-driven external accountability policy*” (see Davis 2006; Kim and Sunderman 2004; LeFloch et al. 2007; Popham 2004 cités dans Lee, 2009, p. 74).

Les États s’engageraient ainsi dans deux voies possibles : celle de la conformité aux attentes — *treat path* — basée sur la coopération entre les États et le niveau fédéral, et celle de la négociation calculatrice – *trick path* – impliquant un rapport de compétition alimenté par les intérêts divergents de ces deux niveaux d’action. La première stratégie (*treat path*) induit une mise en conformité des États vis-à-vis de standards imposés au niveau fédéral, le déploiement et l’alignement d’un ensemble de dispositifs sur cette base. La seconde (*trick path*) se caractérise par un jeu de manipulation où les standards sont négociés par les États, revus à la baisse, et où les performances sont « gonflées » artificiellement, donnant dès lors une image tronquée des performances des élèves.

S’appuyant sur une démarche d’administration de la preuve assez convaincante, Lee suggère que les résultats des États ayant choisi la voie de « l’astuce » (*trick path*) sont ceux qui enregistrent logiquement des gains significatifs de performances, artefacts résultant dès lors des manœuvres mises en place. Les scores des élèves au *testing* national s’envolent, ces derniers étant évalués à l’aune de standards de performance moins exigeants. Les États ayant choisi la voie de la conformité (*treat path*) peinent, quant à eux, à démontrer une amélioration notable de leurs performances du fait de standards trop contraignants, voire inatteignables.

Ce processus de médiation institutionnel, autrement nommé découplage, mettrait donc en partie à mal la stratégie de changement institutionnel au cœur de la politique fédérale. Les conclusions de Lee (2010) doivent pourtant être nuancées dans la mesure où nombre d’auteurs ont par ailleurs mis en exergue la réduction drastique de la marge de manœuvre des acteurs et l’influence des SABP sur les pratiques institutionnelles¹⁸.

Cette étude permet néanmoins d’éclairer les stratégies mises en place par les acteurs dans un contexte de pression à la performance. Ces dernières ainsi que les effets imprévus, ou « pervers », des SABP ont en effet été documentés par plusieurs études à d’autres niveaux

¹⁸ Voir en particulier Coburn (2004) dans le contexte des États-Unis, mais aussi Maroy, Brassard, Mathou, Vaillancourt et Voisin (2015) dans le contexte du Québec.

d'actions. Au niveau des établissements scolaires, ces stratégies se matérialisent dans la falsification des résultats des élèves, l'exclusion des élèves sous performants du *testing*, l'augmentation des taux de redoublements dans les cycles qui précèdent les niveaux dits à enjeux, le placement des élèves sous performants dans des parcours ou des catégories exemptés de *testing* : « maîtrise insuffisante de la langue » – *limited proficiency english*, éducation spéciale (Figlio & Loeb, 2011; Grodsky, Warren, & Kalogrides, 2009; Ladd & Lauen, 2010; Linn, 2000; Mons, 2009; Webb, 2005).

Nombre d'études ont par ailleurs mis en exergue le phénomène maintenant fort connu du *teaching to the test*. Ce phénomène visant à améliorer la réussite des élèves au *testing* induit une focalisation des enseignements sur des objectifs strictement cognitifs, et à court terme, ainsi que le développement de pratiques pédagogiques axées sur la mémorisation rapide. Il entraîne une fragmentation et un rétrécissement du curriculum enseigné, plus particulièrement dans un contexte où la concentration d'élèves sous performants est importante et les performances faibles (Amrein & Berliner, 2002; Darling-Hammond, 2004, 2006; Linn, 2000; Wolf & Janssens, 2007).

Au-delà de pratiques dites stratégiques, la pression liée aux performances et les sanctions attachées aux résultats des élèves et des écoles auraient par ailleurs des effets négatifs à moyen et à long terme, plus particulièrement sur les acteurs œuvrant dans des milieux défavorisés : perte de motivation des enseignants, augmentation du *turn-over* au sein des écoles, difficultés pour ces dernières d'attirer et de retenir des enseignants qualifiés, dégradation de l'image externe de l'école et des professionnels qui y œuvrent, augmentation de la ségrégation académique et scolaire, la somme de ces facteurs participant d'un cercle vicieux et contre-productif (Figlio & Loeb, 2011; Finnigan & Gross, 2007; Mons, 2009; Rustique-Forrester, 2005).

Ces éléments viennent donc questionner les effets des SABP dans leurs formes les plus « dures » et leur capacité à produire l'amélioration escomptée des enseignements et des apprentissages. Ces résultats invitent par ailleurs à user doublement de prudence pour considérer les effets positifs, bien que non consensuels, discutés en amont sur l'amélioration des performances des élèves.

1.12 Conclusion

Au total, si les effets des SABP en termes d'amélioration des performances paraissent relativement ténus, l'équité scolaire semble quant à elle altérée à la baisse, plus particulièrement lorsque les SABP sont assortis d'enjeux forts pour les acteurs. Des dispositifs trop contraignants tendraient ainsi à augmenter la ségrégation scolaire et à pénaliser les publics déjà fragilisés dits « à risque » (Haney, 2000 cité Carnoy & Loeb, 2002; Hanushek & Raymond, 2005).

Il ressort toutefois de cette synthèse de la littérature que les effets des SABP sur l'efficacité et l'équité scolaires ne semblent pas faire l'objet d'un consensus ferme. Les études discutées dans cette section, qui prennent place majoritairement dans le contexte états-unien, mettent en débat des résultats mitigés, tantôt contradictoires, tantôt nuancés par les conclusions d'études subséquentes.

Certains aspects d'ordre empiriques et méthodologiques peuvent toutefois partiellement expliquer les divergences de résultats de ces études. Premièrement, les SABP revêtent des formes organisationnelles variées et s'appuient sur des instruments divers. De fait, il semble difficile de pouvoir statuer sur l'effet (positif ou négatif) de ces derniers de manière générique. En outre, d'un point de vue méthodologique, l'opérationnalisation des SABP, les variables intégrées aux analyses pour caractériser l'efficacité et l'équité scolaires, pour contrôler d'autres facteurs qui influeraient sur les performances des élèves, ainsi que les méthodes d'analyse utilisées témoignent d'une variété non moins importante de paramètres d'analyse clefs qui peuvent influencer sur les résultats produits¹⁹.

Cette littérature et les études qui en constituent le cœur revêtent néanmoins des intérêts multiples. Ces dernières mettent d'une part en exergue les formes diverses que prennent les SABP et, d'autre part, les instruments non moins variés sur lesquels ces derniers reposent. Elles proposent ainsi d'opérationnaliser l'*accountability*, et les SABP, en construisant différents indicateurs et en mobilisant différentes variables pour prendre en compte la diversité de l'objet. Aussi, elles participent à la production d'évidences empiriques quant à leurs effets sur l'efficacité et l'équité scolaires et permettent dès lors de confronter la rhétorique qui entoure le développement des SABP dans les systèmes d'éducation à des résultats factuels.

¹⁹ Ces éléments clefs pour les principales études citées sont synthétisés dans le tableau de synthèse placé en annexe (ANNEXE 1).

1.13 Synthèse de la problématique et perspectives générales de recherche

Ce chapitre nous a permis de mettre tout d'abord en exergue le flou conceptuel qui entoure la notion d'*accountability* pourtant fort usitée dans la littérature. Après avoir discuté ses acceptions multiples et ses formes organisationnelles non moins diversifiées, nous avons mis en perspective le développement des systèmes d'*accountability* basés sur la performance dans les systèmes d'éducation de nombreux pays depuis les années 1990. Nous avons mis en relief le fait que les SABP s'inscrivent dans une conception de la gouvernance renouvelée, fortement influencée par les principes du *New Public Management*, qui participent par ailleurs à la remise en question du rôle de l'État dans la gestion et la production de l'éducation.

Dans ce contexte, les SABP s'appuient sur un ensemble d'instruments visant à réguler les systèmes d'éducation sur base de leurs résultats et sont envisagés comme une réponse politique au déclin postulé de la qualité de l'éducation et des systèmes d'enseignement. Les mécanismes et les instruments sur lesquels ils reposent ainsi que la focalisation induite sur les résultats des élèves devraient ainsi améliorer l'efficacité et l'équité scolaires, notions au cœur des débats qui prennent place à l'échelle nationale et internationale et qui sont alimentés par le développement des grandes enquêtes internationales sur les acquis des élèves.

Nous avons par la suite synthétisé les résultats de la littérature empirique qui vise justement à analyser les effets des SABP sur l'efficacité et l'équité scolaires. Prenant place majoritairement dans le contexte des États-Unis qui se sont engagés massivement sur la voie de l'*accountability*, les études, pourtant nombreuses, suggèrent tout au plus des effets relativement faibles et peu stables sur les performances des élèves. Les scores moyens des élèves augmenteraient légèrement, plus particulièrement en mathématiques et pour les niveaux de scolarité les plus élevés, à la suite de l'introduction de SABP à forts enjeux. En termes d'équité, les effets des SABP apparaissent tendanciellement négatifs. Ces derniers tendraient ainsi à fragiliser les publics et les écoles les plus en difficulté, incapables de faire face à la pression introduite par les SABP. Dans les études discutées toutefois, l'équité est appréhendée majoritairement en termes de différentiel de performances entre élèves issus de différents groupes socioethniques.

Si nous nous sommes plus particulièrement focalisée sur les études états-uniennes qui permettent une certaine comparabilité de leurs résultats, il faut néanmoins rapporter les conclusions de ces

recherches au contexte dans lesquelles ces dernières ont été menées. Les États-Unis se caractérisent en effet par une volonté forte du gouvernement fédéral de réguler le fonctionnement des systèmes d'éducation au moyen d'une approche agressive par les standards, doublée à une régulation par le marché scolaire qui occupe une large place.

Toutefois, des distinctions entre systèmes d'éducation se donnent à voir aux États-Unis dans le choix des instruments qui ont été privilégiés au sein des SABP. On peut dès lors faire l'hypothèse que des outils différenciés caractérisent par ailleurs les SABP tels qu'ils sont mis en œuvre dans d'autres systèmes d'éducation. La littérature mobilisée pour problématiser notre objet nous a en effet permis d'appréhender la multiplicité des outils au cœur des SABP.

Si les SABP s'appuient sur un ensemble d'instruments, de principes et de stratégies de changement relativement communs, ils se développent et s'ancrent néanmoins dans des réalités différenciées qui renvoient aux spécificités des pays et des systèmes d'éducation. Ces spécificités réfèrent aux conditions politiques, idéologiques, matérielles et culturelles qui engagent des conceptions nationales concernant la place de l'État, la valorisation de l'initiative privée, la reconnaissance du pluralisme et la liberté d'enseignement ou encore, des conceptions différenciées du bien éducatif (Maroy, 2008). Ces conceptions influent fortement sur les contextes institutionnels. En cela, des évolutions transversales aux systèmes éducatifs contemporains s'observent différemment en fonction des contextes nationaux et locaux. Des mesures proches peuvent dès lors revêtir des formes d'application particulières et sensiblement différentes (Maroy, 2008).

Toutefois, trop peu d'études se sont, à notre connaissance, intéressées aux effets des SABP dans des contextes hors États-Unis. Si les recherches de Wößmann (2005), Schütz, West et Wößmann (2007) constituent des références en la matière, faisant partie des rares ayant analysé leur impact dans une perspective internationale, tout se passe comme si les SABP mis en œuvre dans d'autres pays du globe faisaient la relative économie d'une évaluation comparative à la lumière de critères d'efficacité et d'équité, pourtant au cœur de la rhétorique qui entoure leur développement.

C'est le cas, nous semble-t-il, de la plupart des systèmes d'éducation d'Europe, mais aussi du Canada. Si certaines études majoritairement qualitatives ont montré que les SABP y reposent *a priori* sur des formes moins « dures » (Dulude & Dupriez, 2014; Maroy & Voisin, 2014; Mons &

Dupriez, 2011), il n'en demeure pas moins que leurs effets en termes d'efficacité et d'équité scolaires doivent, de notre point de vue, être questionnés.

Ainsi, il nous semble que plusieurs constats et questionnements, qui guident les perspectives générales de notre recherche, méritent d'être discutés.

Tout d'abord, la littérature à partir de laquelle nous avons mis en perspective le développement des SABP dans le champ éducatif laisse apparaître un flou conceptuel autour de la notion. Il nous semble donc qu'un effort de conceptualisation doit être réalisé pour saisir et analyser l'objet dans le champ éducatif.

Par ailleurs, si les potentialités supposées des SABP semblent faire l'objet d'un consensus auprès des décideurs politiques, la littérature s'intéressant à leurs effets dans les systèmes d'éducatifs, bien qu'elle apporte certaines évidences empiriques, laissent davantage place, de notre point de vue, à un ensemble de questionnements.

Premièrement, les études situées majoritairement dans le contexte des États-Unis suggèrent un effet différencié des SABP en fonction des instruments sur lesquels ils s'appuient et des types qu'ils revêtent. Ces effets diffèrent par ailleurs selon qu'on s'intéresse à l'efficacité ou l'équité scolaires.

On constate ainsi un manque notable au niveau des recherches quantitatives comparatives internationales. Or, les SABP se développent aussi dans d'autres contextes et s'appuient sur la même rhétorique de l'efficacité et de l'équité scolaires. Les provinces canadiennes s'engagent en effet avec force depuis les années 1990 sur la voie de l'*accountability*. La politique de Gestion Axée sur les Résultats au Québec est ainsi un exemple de la volonté du gouvernement de mettre en place un SABP au sein du système éducatif québécois, les nombreuses réformes mises en œuvre dans le système éducatif Ontarien ou en Alberta en sont un autre (Maroy, Mathou, Vaillancourt & Voisin, 2014; 2015; Chang, Fisher & Rubenson, 2007; Lessard & Desjardins, 2011). En Europe, par ailleurs, les SABP sont mis en œuvre depuis un certains temps dans certains systèmes éducatifs comme celui de l'Angleterre ou de l'Écosse (Broadfoot, 2000; Ozga, 2009; Ozga & Grek, 2010). Il se développent en outre en Suisse (Berhens, 2006; Yerly, 2014), en France (Maroy, Pons, et al., 2017; Pons, 2014), selon des modalités diverses et une instrumentation non moins variées.

Deuxièmement, malgré l'importance accordée aux notions d'efficacité et d'équité scolaires dans la rhétorique politique, ces notions restent peu problématisées. En outre, dans les recherches s'intéressant aux effets des SABP, ces notions sont opérationnalisées selon des indicateurs qui peuvent être qualifiés de restreints. Souvent, différentes conceptions de l'égalité, pourtant à la base de la notion d'équité, ne sont pas questionnées simultanément.

Notre recherche doctorale vise donc à combler ces manques dans la littérature. Au travers d'une étude comparative quantitative internationale, nous souhaitons dépasser le cadre des frontières états-uniennes pour analyser et questionner les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation. Nous intégrons ainsi à l'analyse des systèmes éducatifs peu étudiés jusqu'à lors dans cette perspective : ceux du Canada et de l'Europe.

Par ailleurs, nous tâchons de combler un manque en ce qui a trait à la conceptualisation et à l'opérationnalisation de l'objet. Il s'agit de situer les SABP dans des cadres d'interprétation plus larges, de dépasser l'opposition SABP à forts et à faibles enjeux sur laquelle se base la majorité des recherches quantitatives sur la question, en prenant en compte l'instrumentation au cœur des SABP. Cette dernière s'avère en effet une clef centrale pour différencier d'une part les SABP dans divers contextes, d'autre part pour en analyser les effets.

Aussi, l'analyse ne peut faire l'économie d'un effort de conceptualisation et d'opérationnalisation des notions d'efficacité et d'équité scolaires. Plus particulièrement, il nous semble nécessaire de replacer au cœur de l'analyse différentes figures de l'égalité (et des inégalités, au cœur de la notion d'équité) peu discutées dans les études sur la question.

Il ne s'agit donc pas uniquement de combler un manque dans la littérature et de contribuer à la production d'évidences empiriques sur les effets des SABP sur l'efficacité et l'équité scolaires. Nous souhaitons aussi questionner le développement des SABP dans le champ éducatif. Aussi, s'intéresser aux questions d'efficacité et d'équité scolaires amène de fait à questionner le rôle de l'école, et plus largement de l'éducation.

Notre recherche s'articule ainsi autour du questionnement central suivant : quels sont les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation ?

Cette question centrale appelle des sous questions : comment s'opérationnalisent les SABP dans les systèmes éducatifs d'Europe et du Canada ? Les SABP ont-ils une influence sur les degrés d'efficacité et d'équité scolaires de ces systèmes d'éducation ?

Le cadre général de notre recherche étant posé, le chapitre suivant présente et discute les concepts au cœur de notre étude. Plus particulièrement, notre cadrage théorique s'attachera à discuter et à définir les notions d'*accountability* et de SABP, d'efficacité et d'équité scolaires. Ce travail de conceptualisation nous permettra en outre d'affiner le questionnement au cœur de notre recherche doctorale, de décliner les questions spécifiques et les hypothèses qui sous-tendent notre étude et d'introduire par là même le dispositif méthodologique déployé.

CHAPITRE 2 : CADRAGE THÉORIQUE

Ce chapitre pose le cadrage théorique de notre recherche et discute les concepts centraux qui y sont mobilisés. Le chapitre précédent nous a permis de mettre en perspective l'avènement d'une « nouvelle régulation de l'éducation » et le développement des systèmes d'*accountability* basés sur la performance (SABP) dans le champ éducatif. Nous y avons présenté le questionnement central de notre thèse qui porte sur les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs. Les notions centrales de SABP, d'efficacité et d'équité scolaires seront donc ici discutées, leurs acceptions arrêtées. Il nous semble, en effet, que ces notions ne peuvent faire l'économie d'un effort de définition du fait du flou conceptuel qui les entoure. Par ailleurs, les notions d'efficacité et d'équité scolaires, concepts ombrelles fort mobilisés par la rhétorique politique, doivent être réinterrogées et précisées en même temps que leurs enjeux dans le champ éducatif doivent être éclairés.

Ce chapitre est divisé en deux sections. La première permet de situer conceptuellement la notion de systèmes d'*accountability* basés sur la performance. Nous y ferons un retour sur la notion d'*accountability* déjà discutée dans le chapitre précédent. Nous éclairerons finalement la pertinence de l'entrée d'analyse que nous choisissons pour caractériser les SABP dans les systèmes éducatifs et analyser leurs effets sur l'efficacité et l'équité scolaires.

La seconde examine les notions d'efficacité et d'équité scolaires afin d'en stabiliser les acceptions. Après avoir discuté du concept d'efficacité et ses implications dans le champ éducatif, nous nous pencherons sur le concept d'équité scolaire. Nous discuterons les différentes figures des inégalités au cœur de notre travail de thèse doctorale (inégalité des chances de résultats, des acquis de base et de résultats) à la lumière des travaux en sociologie de l'éducation et de ceux de philosophie politique qui permettent d'éclairer les conceptions de justice qui sous-tendent ces dernières.

SECTION 1 : Régulation de l'éducation et systèmes d'*accountability* basés sur la performance

Cette première partie vise donc à définir et situer conceptuellement les systèmes d'*accountability* basés sur la performance. Un détour par la notion de régulation nous permettra tout d'abord de mettre en perspective les SABP dans un contexte d'évolution des modes de régulation institutionnelle des systèmes éducatifs. Puisque les instruments au cœur des SABP constituent notre entrée analytique pour caractériser ces derniers dans les systèmes d'éducation d'Europe et du Canada, nous mobiliserons et discuterons le concept de mode de régulation institutionnelle. Nous ferons par ailleurs un retour sur la notion d'*accountability* pour stabiliser l'acception des SABP qui sera au cœur de notre recherche. Finalement, nous introduirons la pertinence de notre entrée d'analyse pour caractériser et opérationnaliser les SABP en nous référant à la notion d'instruments d'action publique.

2.1 La régulation des systèmes d'éducation, les apports de la sociologie

Il nous semble fructueux de mobiliser la notion de régulation pour amorcer la conceptualisation de notre objet d'étude. S'intéresser aux SABP dans le champ éducatif revient en effet à porter la focale sur « les modes d'orientation, de coordination et de contrôle de l'action éducative » (Dutercq & Van Zanten, 2011, p. 5) et donc à analyser les modes de régulation institutionnelle qui sont dès lors privilégiés au sein des systèmes d'éducation. Si le concept de régulation réfère dans notre travail à « l'ensemble des arrangements institutionnels et des mécanismes de contrôle et de cadrage des actions promus par une autorité politique reconnue » (Maroy, 2017, p. 734), il a fait l'objet de nombreux travaux dans le champ de la sociologie.

Le concept de régulation est d'abord issu des sciences de l'ingénieur, de la physique et des sciences biologiques pour s'étendre à l'étude des systèmes sociaux. En sociologie, entendue au sens cybernétique et fonctionnaliste, la régulation « tend à assurer l'équilibre d'un processus à l'œuvre ou à rétablir cet équilibre quand “la structure normale est troublée” » (Encyclopédia Universalis citée dans Commaille & Jobert, 1998; Maroy, 2006; 2017). Elle réfère alors aux « mécanismes de rétroaction, aux processus de contrôle et d'échange à travers lesquels un système tend à demeurer en équilibre dynamique, guidé par un objectif de référence » (Maroy, 2006, p. 11). Cette définition appliquée aux processus sociaux se rapproche de la notion de

Social Control « pour désigner le processus d'ensemble qui contribue (...) à assurer le maintien et la permanence de la structure sociale » (l'Ecuyer 1986 cité dans Commaille & Jobert, 1998, p. 19). Cette acception met en relief une dimension essentielle de la régulation : elle participe d'un « processus d'ajustement, de correction des déséquilibres qu'on cherche à réduire et corriger » (Maroy, 2006, p. 14). Elle suppose, par ailleurs, l'existence d'un mécanisme général visant à garantir un certain équilibre du système.

L'introduction des notions de conflit et de dysfonctionnement constitue une alternative fondamentale aux approches fonctionnalistes et consensualistes marquant les sciences sociales dans leur conception de la société. Selon une vision agonistique, en effet, le conflit est considéré comme inhérent à toute société et nécessaire à son évolution. Le fonctionnement des sociétés s'établit dans cette perspective « selon une régulation sociale constituée d'équilibre maintenu ou retrouvé, mais également de tension et de ruptures, de contradictions provoquées par la multiplicité des instances et acteurs sociaux impliqués, par la pluralité des stratégies sociales à l'œuvre » (Commaille & Jobert, 1998 p. 23). Les travaux sur les relations et les conflits dans l'entreprise ont dans ce sens souligné l'importance de la négociation pour construire « des compromis entre rationalités et prétentions contradictoires » (Reynaud, 1980 cité par Commaille & Jobert, 1998, p. 23).

La sociologie des organisations et des relations industrielles s'intéresse quant à elle à l'action collective pour comprendre la production de l'ordre social (Crozier, 1963; Reynaud, 1989; 1999 cités dans Commaille & Jobert, 1998; Maroy, 2006). Elle met en avant les systèmes intermédiaires et les rapports complexes entre différents niveaux d'action (central, local). Elle introduit, par ailleurs, les concepts de pouvoir (Thoenig, 1998), d'acteur, de stratégie au cœur de l'analyse (Muller, 2000, p. 192).

La théorie de la régulation développée par des économistes (Boyer, 1986; Boyer & Saillard, 1995; Boyer, 2004 cités dans Maroy, 2008, p. 14) et la théorie de la régulation sociale travaillée par des sociologues (Reynaud, 1989; 1999; Terssac, 2003 cités dans Maroy, 2008, p. 15) ont par ailleurs tenté d'articuler les notions de régulation et de conflit, en mettant l'accent, d'une part, sur le rôle régulateur des institutions du marché du travail et, d'autre part, sur le fait que la régulation « est un processus pluriel, inachevé, qui procède essentiellement d'un travail de

négociation entre acteurs (...) une action et non une propriété d'un système » (Maroy, 2008, p 15).

S'appuyant sur ces dernières et sur des instruments inspirés de la *New Political Economy* (Jessop, 1995; Boyer, 1997 cité dans Maroy, 2006, p. 16), une perspective constructiviste vient distinguer de manière idéale typique différentes formes de régulation : État, hiérarchie, marché, communauté, réseaux. Elle cherche ainsi à combiner différentes approches de la régulation sociale développée dans la littérature. Elle s'intéresse à la combinaison de différents modes de régulation à différents niveaux d'action, à leur mise en acte négociée au niveau local, à leurs usages contextualisés et à leur évolution « sous l'influence de politiques éducatives nationales, mais aussi en relation avec des évolutions sociales, locales ou globales » (Maroy, 2006, p. 1).

Selon cette approche constructiviste, nous pouvons définir la régulation comme les « processus multiples, contradictoires, conflictuels parfois, d'orientation des conduites des acteurs et de définition des "règles du jeu" dans un système social » (Maroy, 2008, p. 5). Cette perspective permet donc de dépasser une conception de la régulation comme un mécanisme général d'ajustement et une propriété du système qui produirait nécessairement de l'ordre et l'équilibre d'un système social donné. La régulation est ainsi un processus complexe, potentiellement contradictoire qui peut générer désordres et contradictions.

Par ailleurs, la régulation provient de plusieurs sources et prend plusieurs formes (Barroso, 2000; Maroy, 2006; 2008). Institutionnelle et politique, elle réfère à « l'ensemble des arrangements institutionnels et des mécanismes de contrôle et de cadrage des actions promus par une autorité politique reconnue » (Maroy, 2017, p. 734). Elle renvoie en ces termes aux « modes d'action et stratégies » mis en place par la puissance publique, en d'autres termes par l'État, pour coordonner le système et réguler son fonctionnement. Cette régulation formelle appelle cependant d'autres formes de régulation : les régulations autonomes qui interagissent avec les régulations de contrôle promues par les autorités politiques et organisationnelles. Ces régulations autonomes relèvent d'accords multiples, négociés par les acteurs, souvent essentiels à l'adaptation des dispositifs formels aux spécificités locales (Maroy, 2006, p. 12). Elles réfèrent donc aux usages, aux appropriations et aux transformations des dispositifs en fonction des contextes locaux.

Dans le champ éducatif, la régulation institutionnelle, dite de contrôle ou formelle, s'est traditionnellement exercée selon un modèle dit « bureaucratique professionnel » (Barroso, 2000; Demailly, 2001; Maroy, 2006, 2007) fondé sur la cohabitation d'une régulation « étatique, bureaucratique, administrative » et « professionnelle, corporative, pédagogique » (Barroso, 2000, p. 64). Ici, l'État prend en charge la mise en œuvre du service éducatif, selon une forme organisationnelle fondée sur la standardisation et la conformité des règles à l'échelle de l'État nation pour assurer l'égalité de traitement et d'accès pour tous (Maroy, 2006). La régulation est ainsi fondée sur la définition précise des rôles, des fonctions et des compétences de chacun des agents et sur un « ordre hiérarchique des fonctions qui va de pair avec un système d'autorité, c'est-à-dire un système de prise de décision réglementaire légitime, un système de régulation des conflits par les supérieurs hiérarchiques » (Maroy, 2007, p. 10). Ce modèle repose sur la division du travail éducatif (*e. g.* verticale, qui renvoie à l'organisation en termes de niveaux de scolarité, et horizontale, avec le découpage des matières). La dimension professionnelle du modèle bureaucratique professionnel renvoie par ailleurs à l'autonomie professionnelle des agents éducatifs, plus particulièrement celle des enseignants. La régulation de leurs activités (gestion de la classe, travail accompli avec les élèves, méthodes de travail, etc.) est confiée aux mains des enseignants qui disposent dès lors d'une importante autonomie dans la réalisation de leur travail :

« L'enseignant tend à voir son activité s'organiser sur la base d'un modèle proche du modèle professionnel dans la mesure où il se voit confier la responsabilité d'une classe ; le travail qu'il va accomplir avec ses élèves du point de vue du contenu (programme) et de la méthode de travail (pédagogie : formes de transmission, rythme, etc.) relève alors pour l'essentiel de sa propre responsabilité. Il dispose à cet égard d'une grande autonomie, qui lui permet de développer des relations importantes avec ses élèves, qui lui permet également de s'adapter aux différences de niveau, de cursus scolaires antérieurs de ses élèves. C'est dire que selon Bidwell, on tend à se trouver dans une situation où l'activité est organisée de façon proche de celle qui caractérise les activités marquées par ce que les sociologues anglo-saxons appellent un modèle professionnel » (Maroy, 2007, p. 10)

Les formes classiques de la régulation dite formelle ou de contrôle sont néanmoins marquées par des transformations importantes dans un contexte de remise en question de l'État providence et, dans le champ éducatif plus particulièrement, du modèle de l'État éducateur. L'action gouvernementale (Commaille & Jobert, 1998) ne repose plus uniquement sur la loi, les normes et les procédures (d'où leur caractère post-bureaucratique), mais intègre un ensemble de dispositifs qui devraient, d'une part, permettre aux États de reconquérir leur place dans un contexte où un plus grand nombre d'acteurs (publics ou privés) bénéficient d'une influence grandissante dans

les processus décisionnels (au niveau local, national, mais aussi au niveau supranational), d'autre part, soutenir leur légitimité et renforcer leur efficacité (Théret, 1998, p. 100). On se distance alors clairement de l'acception anglophone du terme *regulation* qui fait référence à la régulation par la législation.

La mise en œuvre des systèmes d'*accountability* basés sur la performance s'insère donc dans ce contexte de « modifications substantielles dans la manière dont l'État met en place, fournit, administre et rend compte des services éducatifs publics » (Lessard, 2002, p. 12). Des modes de régulation dits « post-bureaucratiques » sont progressivement introduits. L'art de gouverner et l'action gouvernementale (Commaille & Jobert, 1998) opèrent dès lors au travers d'un ensemble de modes de régulation intégrant un ensemble de dispositifs qui prennent la forme de standards, d'objectifs de performance, d'instruments de contractualisation, d'évaluation²⁰.

Cette régulation systémique ayant trait à « l'ensemble des mécanismes par lesquels un système social se reproduit ou se transforme » (Demailly, 2001, p. 15) appelle par ailleurs une régulation normative qui « concerne le rapport des sujets aux normes et permet l'orientation des actions par des modèles cognitifs et normatifs culturellement et historiquement situés (Demailly, 2001; Maroy, 2008). Ainsi, les transformations à l'œuvre dans le champ éducatif ne s'appréhenderaient pas uniquement en termes systémiques. La recherche d'efficacité (Demailly 2001; Maroy, 2008, 2009), la valorisation de la performance attachée à la mesure des résultats (Duran, 1999 cité dans Maroy, 2008; Ozga, 2009), la visée d'amélioration de la qualité définie en termes de performances des élèves à des tests standardisés portant sur des objectifs cognitifs tendent ainsi à être érigées comme des valeurs fondamentales. On assisterait donc à une transformation du cadre normatif allant de l'obligation de moyens vers une obligation de résultats érigée comme nouveau référent symbolique (Demailly, 2001; Lessard & Meirieu, 2004).

La notion de régulation travaillée dans le champ de la sociologie nous permet donc de situer la mise en œuvre des SABP dans des cadres plus larges d'interprétation qui renvoient finalement au rôle de l'État, à ses moyens d'action, mais aussi aux implications systémiques et normatives des modes de régulation qui sont introduits pour réguler les systèmes d'éducation par leurs résultats.

²⁰ Voir sous-section 1.4 Systèmes d'*accountability* basés sur la performance, instruments privilégiés.

2.2 Modes de régulation institutionnelle et instruments d'action publique

Nous avons déjà employé nombre de fois le terme de modes de régulation institutionnelle. Si nous l'avons défini de manière succincte, rappelons que cette notion fait référence à la façon « dont les autorités publiques exercent la coordination, le contrôle et l'influence sur le secteur éducatif, orientant au travers de normes, d'injonctions et de contraintes le contexte de l'action des acteurs sociaux et leurs résultats » (Barroso, 2004, p. 92 cité dans Maroy, 2006). Les modes de régulation institutionnelle renvoient en ces termes à divers instruments — les normes, les lois, la contractualisation, les dispositifs d'évaluation (etc.).

Plusieurs travaux se sont intéressés à la question des instruments comme entrée analytique pour étudier l'action publique et les « modalités concrètes de l'exercice du pouvoir ». Les travaux de Foucault se sont en effet intéressés à « l'instrumentation en tant qu'activité centrale dans l'art de gouverner » (Senellart, 1995 cité dans Lascoumes & Le Galès, 2004, p. 17), ceux de Weber problématisent les instruments d'action publique en tant que « technique de domination » (Lascoumes & Le Galès, 2004, p. 16).

Plus récemment, parallèlement aux transformations des modes d'intervention de l'État, les analyses portant sur l'instrumentation de l'action publique ont connu de forts développements. Elles s'intéressent d'une part aux techniques d'intervention de l'État et à la recherche d'instruments plus efficaces dans ce que Lascoumes et Simard nomment une « perspective classique de *policy analysis* » (Lascoumes & Simard, 2011, p. 9), d'autre part à « la diversification constante des instruments et aux problèmes de coordination qui en découlent » (Lascoumes & Simard, 2011, p. 11), enfin aux usages par les acteurs et aux effets qui leur sont liés.

Des travaux récents se sont par ailleurs intéressés à la dimension normative des instruments d'action publique et font donc écho à la régulation normative discutée précédemment. Les instruments d'action publique y sont considérés comme des dispositifs techniques liés à des finalités d'action qui orientent les comportements des acteurs et structurent par ailleurs les relations entre les individus. Les instruments d'action publique sont entendus comme des institutions (au sens sociologique du terme), ils réfèrent en ces termes à « un ensemble plus ou moins coordonné de règles et de procédures qui gouverne les interactions et les comportements

des acteurs et des institutions » (North, 1990 cité dans Lascoumes & Legalès, 2004, p. 15).
Lascoumes et Simard (2011) en proposent la définition suivante :

« un dispositif normatif, à la fois technique et social, ayant une vocation générique (applicable à un ensemble de situations) et porteur d'une conception concrète du rapport gouvernant/gouverné fondé sur une conception spécifique de la régulation » (Lascoumes & Simard, 2011, p. 19).

Cette acception est intéressante dans la mesure où elle remet en question la neutralité supposée des instruments (*value free*). Les instruments d'action publique et les modes de régulation institutionnelle ne sont pas de simples dispositifs techniques, mais sont aussi porteurs d'un sens et d'une certaine conception de la situation, du problème, de la question qui justifient leur conception et leur mise en œuvre, en d'autres termes, d'une certaine logique de régulation.

Pour Lascoumes et Simard (2011) l'analyse de l'instrumentation au cœur de l'action publique permet finalement d'appréhender cette dernière en action. Une entrée par les instruments devrait dès lors nous permettre d'éclairer les modalités de contrôle privilégiées pour piloter les systèmes d'éducation sur base de leurs résultats, et donc, de mettre au jour les instruments centraux sur lesquels reposent les SABP dans divers contextes.

Avant de discuter certains travaux qui ont privilégié l'instrumentation pour analyser et typifier les SABP dans le champ éducatif, il nous semble au préalable nécessaire d'effectuer un retour sur la notion d'*accountability* à la lumière des concepts que nous venons de discuter.

2.3 Retour sur la notion d'*accountability*

Nous avons, dans le premier chapitre de notre thèse, discuté des acceptions de la notion d'*accountability*, concept polysémique qui fait droit à des interprétations diverses résumées par Veselý :

“an unspecified umbrella term, often normatively (e.g., in different strategic documents of governments) (...) In this sense, accountability is used as a synonym for many desirable, yet loosely defined political goals, such as good governance, transparency, or democracy (Bovens, 2010, p. 946)

The term is sometimes used in a broad sense, as different types of mechanisms (both external and internal) for controlling and ensuring quality in public institutions and for making them sensitive to the citizens' demands (such mechanisms may include the market) (such mechanisms may include the market, transparency, trust etc.). Unfortunately, both the first and the second meanings of the accountability are too general and ambiguous

Third, the term “accountability” is used in the narrower sense, as a kind of social relation between two actors in which one actor (an individual or an institution) answers for his actions to the other actor who has the right to judge those actions, demands corrections and, if necessary, imposes sanctions” (Vesely, 2013, pp. 5-6)

Nous avons pour notre part choisi une définition large de la notion d'*accountability* qui nous permet de dépasser l'acception de l'*accountability* comme mécanisme, travaillée par Bovens (2007, 2010), et la focalisation sur la reddition de comptes, pour privilégier une acception qui permet de prendre en compte un ensemble d'instruments au cœur des SABP.

Cette acception se rapproche ainsi de la notion de régulation par les résultats mobilisée dans la littérature francophone, qui permet d'intégrer à l'analyse l'ensemble des modes de régulation institutionnelle visant à piloter le système par ses résultats.

Les SABP font ainsi référence dans notre travail à une expression particulière de la régulation institutionnelle ou de contrôle mise en place par une autorité publique et étatique. Ils constituent en ces termes une forme de régulation formelle qui s'appuie sur et convoque, à l'échelle des systèmes d'éducation, un ensemble d'instruments pour contrôler, orienter et coordonner les actions des acteurs.

La définition empruntée à Rhoten et collègues (2003), à laquelle nous avons déjà fait référence, permet dès lors de situer les SABP dans une double perspective instrumentale associée aux finalités poursuivies. Nous la reprenons ci-dessous à toutes fins utiles:

“The use of tests procedure, methods, or series of tasks to measure what is taught and learned (Leithwood, Edge and Jantzi, 1999) an arrangement whereby ‘an account must be given to some authority, as an indication of compliance with defined standards, and as demonstrated by improvement on baseline or performance measure’ as determined by some form of assessment” (Rhoten et al., 2003, p. 7)

Les implications des SABP sont en outre soulignées par Furhman²¹ (1999) : 1) une focalisation sur les résultats des élèves et sur les performances des acteurs, qui font dès lors l'objet d'un contrôle accru ; 2) la définition des établissements scolaires comme niveau d'action et « unité d'amélioration » ciblés ; 3) la mise en place de stratégies d'amélioration des écoles ; 4) le développement de nouvelles modalités d'inspection et d'évaluation des établissements scolaires ; 5) la définition de catégories d'imputabilité pour les écoles ; 6) la publication des résultats des

²¹ Ces éléments caractérisent la nouvelle *accountability* selon Furhman (1999).

élèves et des écoles ; 7) la mise en œuvre d'un ensemble de conséquences attachées aux performances des acteurs.

Dans notre recherche, une importance centrale est donc accordée aux instruments pour appréhender les SABP dans les systèmes éducatifs. Dès lors, une série de travaux qui analysent les SABP à partir de leurs instruments sont fort pertinents pour notre propos. La discussion de ces travaux et de leurs apports, dans la sous-section suivante, permettra de clore la première partie de notre cadrage théorique.

2.4 Caractériser les SABP par leurs instruments

Plusieurs auteurs ont en effet choisi l'instrumentation comme entrée d'analyse pour différencier les SABP dans les systèmes d'éducation. Certains travaux aboutissent ainsi à la construction de typologies permettant de différencier les SABP selon des types empiriques distincts à l'instar de certaines études états-uniennes discutées en amont²² (Carnoy & Loeb, 2002 ; Amrein & Berliner, 2002 ; Lee & Wong, 2004). D'autres travaux ont construit quant à eux des typologies théoriques proposant des catégories analytiques permettant d'apprécier la diversité des instruments au cœur des SABP et de questionner les logiques qui les sous-tendent.

Dans la première série de travaux, les études statistiques discutées au chapitre précédent proposent d'opérationnaliser l'*accountability* en prenant en compte les instruments qui en constituent le socle. Dans cette veine, les travaux d'Amrein et Berliner (2002), de Carnoy et Loeb (2002), mais aussi de Lee et Wong (2004) sont, de notre point de vue, les plus aboutis. Leurs démarches de typification empirique aboutissent à la construction d'indicateurs composites visant à différencier les systèmes éducatifs états-unis. Les auteurs prennent en compte l'étendue des instruments sur lesquels reposent les SABP et opèrent une classification à partir, principalement, des enjeux associés aux résultats des élèves pour différents acteurs. Ils permettent donc de distinguer, dans le contexte des États-Unis, les SABP à forts enjeux qui intègrent des conséquences sociales et matérielles importantes pour les acteurs, et les SABP à faibles enjeux qui reposent sur une gamme d'instruments moins étendue et où les performances des élèves ne sont pas nécessairement liées à des enjeux formels pour les acteurs.

²² Voir SECTION 2 : Effets des systèmes d'*accountability* basés sur la performance sur l'efficacité et l'équité scolaires, revue de la littérature

Ces travaux renvoient ainsi à la distinction classique qui émerge dans la littérature entre systèmes d'*accountability* à forts et à faibles enjeux, ou d'*accountability* dure versus douce (Dulude & Dupriez, 2014; Mons, 2009; Mons & Dupriez, 2011). Cette distinction repose d'une part, et très centralement, sur les enjeux associés aux performances des élèves. D'autre part, ces deux modèles reposent finalement sur des théories de l'action et sur des conceptions de l'acteur sensiblement différentes telles que développées dans les travaux de Maroy et Voisin (2014), que nous présentons ci-dessous et qui s'appuient en partie sur ceux de Mons (2009), de Dupriez et Mons (2011).

Dans la seconde série de travaux, Maroy et Voisin (2014) mobilisent en effet aussi l'instrumentation comme clef pour analyser les SABP. La visée de ces travaux est toutefois sensiblement différente. Il ne s'agit pas de classifier les systèmes éducatifs à partir des instruments mobilisés dans les SABP, mais de proposer une typologie théorique permettant, sur base de l'instrumentation au cœur des SABP, de mettre au jour différentes logiques de régulation. Leur typologie offre dès lors des repères analytiques pour « penser » les SABP dans les systèmes d'éducation à la lumière de types théoriques de régulation.

Ces auteurs distinguent ainsi de manière idéale typique (Schnapper, 1999) plusieurs logiques de régulation. Ces dernières se distinguent par la gamme et l'étendue des instruments mobilisés, par le degré d'alignement et de couplage entre ces différents outils et entre niveaux d'action, enfin du point de vue de la théorie de l'action et des conceptions des acteurs incorporées dans les instruments qui sont privilégiés. Selon ces auteurs, en effet, les logiques de régulation au cœur des SABP reposent sur des conceptions différentes des individus et de ce qui est censé les « faire bouger ». Ceci implique dès lors le recours à des instruments distincts qui devraient susciter ou contraindre le changement. Ils distinguent ainsi quatre logiques de régulation : 1) une logique de régulation par une responsabilisation « douce », 2) une logique de régulation par une reddition de comptes et une responsabilisation « réflexives », 3) une logique de régulation par une reddition de comptes néo-bureaucratique sur les résultats, 4) une logique de régulation par une reddition de comptes « dure » sur les résultats.

La logique de « régulation par une responsabilisation douce » (Maroy & Voisin, 2014) repose ainsi sur une faible gamme d'instruments et sur un faible degré d'alignement entre outils et niveaux d'action. Dans cette logique qui fait place à une conception « réflexive » des individus,

l'accent est mis sur la responsabilisation des acteurs quant aux effets de leurs pratiques et aux résultats qu'elles produisent. Les individus y sont considérés comme capables de s'engager de leur propre chef dans un processus de réflexion sur leurs pratiques lorsqu'ils sont confrontés aux résultats de leurs élèves, ce que Thélot nomme l'« effet miroir ». L'accent est donc davantage mis sur des dispositifs d'accompagnement permettant l'intériorisation de normes que sur des dispositifs externes et contraignants pour susciter le changement. Ce processus de responsabilisation et d'intériorisation de normes en lien avec les SABP est en outre explicité dans les travaux de Kogan (1988) :

“A teacher develops standards of appropriate behaviour and of teaching performance as a product of both innate moral propensities and through education and socialisation which are external forms of influence. Once within role, the teacher finds that his internalised norms are modified or strengthened by the expectation placed on him by the school and employment system” (Kogan, 1988, p. 31)

La logique de régulation par une reddition de comptes « dure » sur les résultats repose, quant à elle, sur une large gamme d'instruments fortement alignés les uns aux autres et sur des conséquences matérielles et sociales importantes pour les acteurs. Elle repose sur une conception « utilitariste » des acteurs. Sensibles aux pressions externes, ces derniers devraient mettre en place les correctifs nécessaires et adopter des pratiques plus efficaces, lorsqu'ils sont encadrés et confrontés à un ensemble de dispositifs contraignants visant à cadrer et orienter leurs actions.

Les deux autres logiques reposent finalement sur des variantes entre gammes plus ou moins étendues d'outils et sur des usages intensifs des résultats des élèves à différents niveaux d'action pour responsabiliser les acteurs ou les rendre imputables des performances de leurs élèves. Ainsi dans une logique de régulation par une responsabilisation et une reddition de comptes réflexives » ce sont à la fois des instances d'accompagnement des professionnels qui sont privilégiées, mais aussi des dispositifs contraignants permettant d'encadrer le processus de réflexivité et de l'orienter à la lumière de standards de performances et de pratiques établis. Dans cette logique, l'acteur n'est pas uniquement considéré comme stratège, il a aussi une dimension réflexive et institutionnelle (qui réfère à des « schèmes cognitifs, des habitudes acquises, des modèles normatifs » qui orientent son comportement [Maroy & Voisin, 2014, p. 39]).

Enfin, dans une logique de régulation par une reddition de comptes néo-bureaucratique sur les résultats, l'accent est plutôt mis sur le suivi des résultats des élèves et des écoles ainsi que sur des

procédures de reddition de comptes selon des modalités bureaucratiques et hiérarchiques. À l'inverse de la logique précédente, la conception des acteurs qui prévaut ici est avant tout une conception utilitariste de ces derniers.

Les travaux de Maroy et Voisin (2014), en distinguant différentes logiques de régulation, ont ainsi le mérite de mettre l'accent d'une part sur l'instrumentation au cœur des SABP, d'autre part, sur les théories de la régulation mises en acte au travers de différents outils. Cette construction idéale typique renvoie toutefois à des modèles théoriques « purs » qui sont travaillés et hybridés au sein des systèmes d'éducation. Si elle offre un cadre analytique pertinent, elle devrait néanmoins faire l'objet de confrontations empiriques pour valider et complexifier les cadres d'analyses proposés.

Finalement, Harris et Herrington (2006) particularisent deux types d'*accountability* qui renvoient aussi à deux idéaux types de régulation opposant une régulation par le marché à une régulation par l'autorité gouvernementale. La notion de *government based accountability* renvoie aux efforts entrepris par le gouvernement pour mesurer les résultats et les performances des élèves sur la base des évaluations et des tests standardisés institués à l'échelle des systèmes d'éducation. Elle réfère aux pressions exercées par les autorités centrales pour mettre en œuvre divers instruments dans une visée d'amélioration des performances des élèves et des établissements scolaires liée à un système d'incitatifs et de sanctions.

La notion de *market based accountability* (Harris & Herrington, 2006) renvoie, quant à elle, à la place centrale du marché éducatif autour duquel s'organisent les mécanismes d'imputabilité et de reddition de comptes. Cette forme organisationnelle et les incitatifs qui s'y attachent renvoient à l'*accountability* de marché discutée au premier chapitre de notre thèse, nous n'y revenons donc pas. Précisons toutefois que, dans les faits, les deux types d'*accountability* discutés par Harris et Herrington (2006) tendent à se juxtaposer dans nombre de systèmes d'éducation.

Au total, ces travaux mettent premièrement en exergue la diversité des instruments au cœur des SABP. D'autre part, ils permettent d'éclairer des logiques de régulation différenciées au cœur des SABP portées dès lors, et mises en acte par les instruments qui en constituent le socle. Ils viennent, en outre, appuyer la pertinence de choisir comme entrée d'analyse l'instrumentation pour caractériser les SABP dans divers contextes.

2.5 Conclusion

Nous avons souhaité, dans cette première section de notre cadrage théorique, situer conceptuellement les SABP en ayant recours au concept de régulation et à ses développements dans le champ de la sociologie. Cette notion renvoie en effet à des acceptions différenciées en fonction de la perspective théorique au travers de laquelle elle est appréhendée.

Si nous avons adopté une perspective constructiviste pour stabiliser l'acception de cette notion, les approches discutées dans cette section éclairent deux aspects de la régulation qui nous semblent fondamentaux. Le premier a trait à la visée de coordination, de production de l'ordre et de maintien d'un certain équilibre ; le second relève des processus complexes de médiations, d'appropriation des modes de régulation mis en place par la puissance publique pour réguler et coordonner le système.

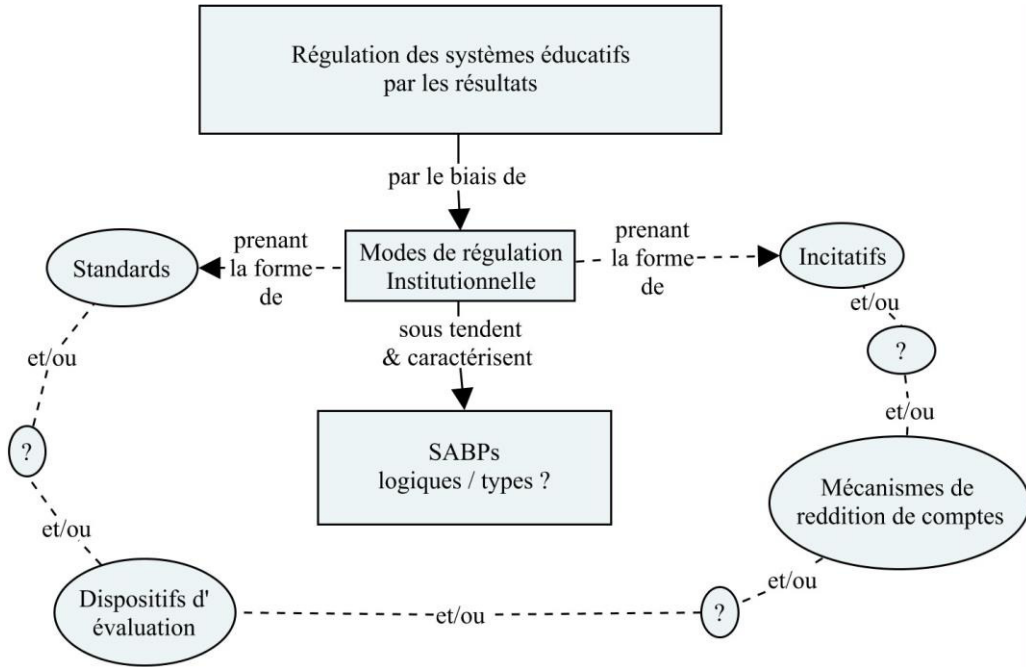
Si le concept de régulation et l'intérêt qu'il suscite dans le champ des sciences sociales s'inscrit dans un contexte de remise en question idéologique, théorique et réelle du « mythe » de l'État tout puissant, omnipotent et omniprésent (Muller, 2000), l'analyse de la littérature montre que, dans le champ éducatif en particulier, plus qu'un affaiblissement de l'État, il s'agirait davantage d'une transformation de ses modalités d'intervention et de son rôle dans la gestion et la production de l'éducation. Le développement des SABP s'inscrit en effet dans le contexte de l'introduction de nouveaux modes de régulation institutionnelle qui remettent en question le modèle de régulation sur lequel ont été construits la plupart des systèmes d'enseignement.

Ces modes de régulation post-bureaucratiques (Maroy, 2006) et l'instrumentation au cœur de l'action publique ont ainsi fait l'objet de développements importants dans la littérature. Dans le champ éducatif en particulier, cette entrée d'analyse a permis à certains auteurs de caractériser et de différencier les SABP au sein des systèmes éducatifs. Malgré certaines limites, cette littérature s'avère un ancrage pertinent pour notre travail de recherche doctorale. Elle justifie en effet d'une part le recours à une acception large de l'*accountability*, d'autre part l'intérêt et la pertinence de caractériser les SABP au travers des instruments qui en constituent le socle, pour pouvoir, par la suite, en analyser les effets sur l'efficacité et l'équité scolaires.

Au total, cette section nous a permis de stabiliser et de discuter de la notion de SABP au travers de différentes notions clefs. Avant de définir les concepts clefs d'efficacité et d'équité scolaires

dans la section suivante, nous proposons, à titre de synthèse, de situer (Figure 2) les SABP au travers des principaux concepts que nous venons d'aborder.

Figure 2 : Schéma conceptuel — régulation des systèmes éducatifs par les résultats et systèmes d'*accountability* basés sur la performance (SABP)



SECTION 2 : Efficacité et équité scolaires

Puisque notre recherche vise à analyser les effets des SABP sur l'efficacité et l'équité scolaires, cette partie précise les acceptions de ces deux concepts clefs. Nous traiterons d'abord de la notion d'efficacité puis nous nous pencherons sur le concept d'équité. Ce dernier concept nécessite de s'intéresser à différentes figures de l'égalité qui sous-tendent les principes normatifs au cœur de la définition, dès lors non univoque, de l'équité scolaire. Dans les sous-sections dédiées à la notion d'équité, nous convoquons la littérature en sociologie de l'éducation, mais également en philosophie politique. Plus particulièrement, nous discutons différentes figures des inégalités et les théories de justice qui s'y rapportent, puisque ces dernières permettent de « penser » la justice scolaire.

2.6 De l'efficacité

L'efficacité est définie par Orivel (1999 cité dans Mons, 2017, p. 306) comme « la capacité d'un système à produire les résultats attendus ». Ce concept multidimensionnel, au cœur de nombreux débats politiques et scientifiques, est donc, pour reprendre les propos de Sall et De Ketele (1997), « de l'ordre de la visée ». Il met en lien les objectifs fixés avec les résultats obtenus. Rapporté aux systèmes d'éducation, il introduit la question de la capacité de ces derniers à atteindre les objectifs établis. Il implique donc d'avoir une idée relativement claire des objectifs poursuivis pour pouvoir y rapporter les résultats obtenus. Or, la question des objectifs assignés aux systèmes d'éducation et d'enseignement appelle des réponses qui sont loin d'être univoques, perpétuellement soumises au débat. Cette question est, en effet, éminemment politique. Elle est aussi stratégique, car elle interroge les finalités de l'éducation autant qu'elle devrait guider l'action éducative.

Trois fonctions sont traditionnellement attribuées à l'école moderne et aux systèmes scolaires dans leurs formes relativement récentes, celles de formation, d'éducation et de socialisation (Thélot & Vallet, 2000). Elles sont déclinées par Dubet et Martucelli (1996), reprises par Crahay (2012) en termes d'éducation, de socialisation et de distribution. La fonction éducative est, selon Crahay, « liée au projet de développement de la personne qui, capable de réflexivité, s'autodétermine et s'autorégule en fonction de ses analyses rationnelles des situations qu'elle est amenée à affronter » (Crahay, 2013, p. 24). La fonction de socialisation réfère à l'intégration de

la culture, des normes et des valeurs inhérentes à la société dans laquelle les individus s'insèrent (Crahay, 2013). L'école doit aussi permettre l'acquisition de savoirs et de connaissances, d'habiletés et de capacités des individus à participer à la vie civile. Enfin, la fonction de distribution renvoie à l'attribution de qualifications scolaires et professionnelles « possédant une certaine utilité sociale » (Dubet & Martucelli, 1996 cités dans Crahay, 2013, p. 25).

L'idée n'est pas ici de discuter de chacune de ces fonctions qui ont fait et font encore aujourd'hui l'objet de débats sur la scène publique, de considérations théoriques et d'interprétations multiples dans la littérature. Il s'agit plutôt de mettre en exergue le fait que chacune d'entre elles peut constituer un objectif à la lumière duquel évaluer les systèmes d'éducation en termes d'efficacité. Il faut néanmoins admettre que, dans le contexte actuel de mutation des sociétés et d'évolution des systèmes éducatifs, les finalités de l'institution scolaire sont fortement questionnées. L'enjeu est de taille puisqu'il est admis de longue date que l'éducation est un facteur de productivité et de transformation culturelle et sociale (Cussó & D'Amico, 2005).

Il faut par ailleurs mettre en perspective les préoccupations croissantes quant à l'efficacité des systèmes d'éducation à la lumière de certains facteurs. Dès les années 1950-1960, qui sonnent le début d'une ère d'expansion économique et la massification de l'enseignement, une préoccupation autour du maintien de la croissance économique impulse aux États-Unis le développement de recherches en économie de l'éducation, ayant pour objet la mesure des investissements éducatifs. Les travaux de Schultz et Becker (cités dans Normand, 2006, p. 34) s'intéressent à conjuguer des analyses coût-efficacité à la question de la rentabilité des systèmes éducatifs. Le lien entre l'éducation, la productivité du travail et la valeur ajoutée, alors au centre des préoccupations des économistes, influencent les recherches en éducation dans l'idée que des « niveaux élevés d'éducation de la population devaient permettre d'augmenter la productivité et de stimuler la croissance » (Normand, 2006, p. 55).

Mons (2004) illustre le lien, au cœur de la théorie du capital humain, entre efficacité économique et éducation :

« Dans la lignée de la théorie du capital humain, l'éducation est non plus considérée comme un coût, mais comme un investissement dont va résulter un capital. L'agrégation des compétences individuelles va refléter partiellement la force économique des pays. Les évaluations internationales des élèves sont ainsi devenues des enjeux politiques importants moins parce qu'elles reflètent la qualité des systèmes éducatifs que parce qu'elles permettent aux différents États de se mesurer à travers leurs potentiels économiques futurs.

Les écoliers d'aujourd'hui sont les travailleurs de demain. Leurs compétences scolaires d'aujourd'hui reflètent leur potentiel économique de demain » (Mons, 2004, p. 60)

Loin de disparaître, les considérations autour de l'efficacité et de l'efficience des systèmes d'éducation, concepts directement issus du vocable économiste, s'accroissent à la fin des années 1980 dans un contexte de crise économique et de recherche de compétitivité des nations sur le marché globalisé de l'économie de la connaissance.

Elles doivent par ailleurs être situées dans le cadre du développement d'instruments et de dispositifs méthodologiques de plus en plus sophistiqués, et à grande échelle - du *testing* au niveau national, des grandes enquêtes internationales sur les acquis des élèves au niveau international, mais aussi de l'essor de la mesure en éducation et des travaux sur l'école efficace dans les années 1980 (Normand, 2006).

Ces préoccupations d'efficacité portent à la fois sur l'efficacité externe et l'efficacité interne des systèmes d'éducation. Si l'efficacité externe amène à « mesurer les effets de la formation dans la vie en dehors de l'école » (Paul, 2005, p. 31), elle réfère selon Mons (2017) « à la relation entre le système éducatif et l'ensemble de la société, à sa capacité à atteindre les objectifs sociaux, économiques et politiques qui lui sont assignés par la société » (Mons, 2017, p. 306). Ainsi, pour Sall et De Ketele :

« Évaluer l'efficacité externe d'un système éducatif ou d'une institution pédagogique, c'est se demander si les individus qui en sortent sont non seulement socialement et économiquement utiles (ou productifs), mais aussi s'ils sont en mesure de développer leur personnalité dans les différentes dimensions (cognitive, affective, psychomotrice, relationnelle ou artistique). L'efficacité externe doit donc tenir compte des objectifs de la société, des besoins du marché du travail et des aspirations individuelles » (Sall & De Ketele, 1997, p. 125).

L'efficacité interne renvoie quant à elle « au fonctionnement du système éducatif en lui-même, à sa capacité à faire en sorte que les objectifs strictement pédagogiques ou scolaires soient atteints » (Mons, 2017, p. 306). Comme pointé par Paul (2005, p. 40), « parler d'efficacité interne dans le cadre du système éducatif conduit à le considérer comme un système de production, particulier certes, mais auquel on pourrait appliquer les mêmes outils que ceux que l'on utilise pour la production des autres biens et services²³ ».

²³ Sur la question de l'éducation considérée comme un service, l'article de C. Maroy (2009) est assez éclairant. Nous invitons le lecteur à le consulter.

Au niveau des systèmes d'éducation, la mesure de l'efficacité interne impose de se doter de différents outils, qui peuvent être mis en perspective par la référence à la définition de l'efficacité proposée par Orivel (1999 cité dans Mons, 2017), liant résultats obtenus et objectifs fixés. Rapportée au développement de la régulation par les résultats des systèmes d'éducation, et à la mise en œuvre des SABP, cette acception éclaire dès lors l'importance accordée à la fixation de standards académiques et de performance²⁴. Il s'agit donc, pour les pouvoirs publics, de définir les objectifs à atteindre, en termes d'une part de connaissances et de compétences que les élèves doivent acquérir, d'autre part, de niveaux de performances attendus. Par ailleurs, les systèmes d'éducation doivent se doter d'outils de mesure permettant d'évaluer l'atteinte des objectifs fixés. L'évaluation standardisée devient donc dès lors, comme nous l'avons vu, un instrument privilégié permettant de contrôler les résultats du système à la lumière des grandes orientations fixées au niveau central.

S'agissant des recherches qui ambitionnent de mesurer l'efficacité des systèmes scolaires, l'enjeu est de mettre au jour les facteurs qui participent à l'augmentation des performances des élèves et des écoles. Elles souhaitent analyser dans quelle mesure l'organisation des systèmes éducatifs et les modes de régulation institutionnelle contribuent à l'amélioration de l'efficacité scolaire. Il s'agit donc finalement d'observer la capacité de plusieurs facteurs « à élever le niveau moyen » des élèves (Attali & Bressoux, 2002).

Ici, la posture comparatiste, « à la base du concept d'efficacité » (Mons, 2017, p. 306), s'impose progressivement comme une « nécessité » car elle permet d'apprécier la variation de certains facteurs difficilement observables dans les contextes nationaux. Ce postulat permet d'éclairer l'intérêt qui est accordé aux grandes enquêtes internationales sur les acquis des élèves, qui tâchent justement de mettre en lien certains facteurs structurels et organisationnels avec l'efficacité scolaire.

Le développement d'un champ de recherche, récent, mais prolifique, qui s'intéresse à l'impact des SABP sur les performances scolaires s'inscrit justement dans cette préoccupation d'efficacité des systèmes d'éducation. L'objectif est, comme nous l'avons vu²⁵, de mesurer la contribution de divers instruments ou types de SABP à l'augmentation des résultats des élèves. C'est dans la

²⁴ Voir sous-section 1.4 Systèmes d'accountability basés sur la performance, instruments privilégiés.

²⁵ Voir SECTION 2 : Effets des systèmes d'accountability basés sur la performance sur l'efficacité et l'équité scolaires, revue de la littérature

lignée de ces études que s'inscrit notre recherche doctorale. Si nous souhaitons analyser l'effet des SABP et de leurs instruments sur le degré d'efficacité interne des systèmes d'éducation, ce dernier renvoie dans notre travail au niveau moyen, en d'autres termes aux performances moyennes, des élèves dans les systèmes d'éducation.

Par ailleurs la question des effets des SABP sur l'efficacité scolaire est articulée, dans notre recherche, à celle de leur impact sur les inégalités de performances entre élèves. Nous nous rangeons ainsi derrière le point de vue de Mons qui précise que la mesure de l'efficacité doit aussi être accompagnée de la mesure des « inégalités selon différents critères » (Mons, 2017). Il s'agit donc d'observer d'une part dans quelle mesure les SABP participent effectivement de l'augmentation du niveau (moyen) de performances scolaires, d'autre part, d'appréhender leur contribution à l'équité scolaire. Nous proposons dans les sous-sections suivantes une discussion autour des notions d'équité et d'inégalités pour stabiliser ces concepts qui sont interrogés dans notre étude.

2.7 De l'équité et des inégalités

L'équité est définie par Psacharopoulos et Woodhall (1988, p. 259 cités dans Sall & De Ketele 1997) comme la « façon dont les coûts et les avantages de l'investissement sont distribués parmi les différents groupes de la société ». Cette définition, *a priori* peu satisfaisante, peut-être étayée en référence aux travaux de Hutmacher, Cochrane et Bottani (2001) qui précisent que le concept d'équité a trait à la question « normative et éthique de l'attribution/acquisition juste de ressources, avantage et désavantages » (Hutmacher, Cochrane & Bottani, 2001, p. 10). Ce concept renvoie ainsi à des questions de justice sociale et, dans le champ éducatif, de justice scolaire. Il est ainsi porteur d'une « conception particulière du rôle et du fonctionnement de l'école » (Mons, 2004, p. 20).

L'équité implique ainsi un ensemble de « valeurs, de critères et de principes engageant un jugement en termes normatifs » (Hutmacher et al., 2001, p. 11). Elle pose la question des inégalités qui peuvent être considérées comme « tolérables », ou à l'inverse, inacceptables, et amène donc à interroger différentes figures de l'égalité (et donc des inégalités) au sein des systèmes d'éducation : elle « juge l'injustice d'une inégalité » (Meuret, 2017, p. 312). Les notions d'équité et d'égalité sont ainsi étroitement liées, bien qu'elles s'avèrent distinctes tel qu'illustré par Mingat (1988) :

« Il semble que le terme d'égalité (ou d'inégalité) ait plutôt une connotation factuelle, alors que le terme d'équité, qui entretient des relations évidentes avec celui d'égalité, incorpore en outre des jugements de valeurs et des conceptions morales » (Mingat, 1988, p. 93)

Si selon Duru-Bellat (Duru-Bellat, 2002), la notion d'équité scolaire implique que l'école doive compenser les inégalités injustes et prévoir un traitement compensatoire de ces dernières, il faut préciser que les préoccupations d'équité des systèmes d'éducation, et le débat autour de la question des inégalités scolaires, sont relativement récents.

« Les individus, si postulés égaux qu'ils soient, sont de fait dans des situations inégales. La notion d'équité suppose qu'on ne conçoive l'égalité des droits qu'en fonction de la situation particulière de chacun (...). Une égalité par la loi doit être visée grâce à des traitements juridiques spécifiques, de type discrimination positive, adaptée aux situations concrètes de chacun » (Duru-Bellat, 2002, p.17).

Ainsi, selon Verhoeven et Dupriez (2008), et Mons (2004), la notion d'égalité est au cœur du projet de l'école moderne.

« La notion d'égalité est partie intégrante du projet scolaire moderne et apparaît comme un des critères majeurs à l'aune duquel le monde scolaire est à la fois pensé, construit et interrogé » (Verhoeven & Dupriez, 2008, p. 95).

L'école moderne marque l'avènement de l'obligation scolaire dans la majorité des sociétés occidentales démocratiques où tous les enfants, ou presque, sont aujourd'hui scolarisés de 6 à 12 ans, en moyenne (Crahay, 2012). Il aura en effet fallu attendre le XIX^e siècle pour que le droit à l'instruction et à l'éducation émergent dans un contexte où, jusqu'ici, le fait que seule une partie de la population, majoritairement issue des classes supérieures, reçoive instruction et éducation ne faisait pas l'objet de contestation (Verhoeven & Dupriez, 2008).

C'est ainsi avec l'école moderne qu'est apparue la notion d'inégalité scolaire (Mons, 2004) et le débat sans cesse réactualisé sur les inégalités au sein des systèmes d'éducation qui s'articule autour de diverses questions clefs : quels sont le rôle et la responsabilité des systèmes éducatifs dans la « fabrication » des inégalités scolaires ou la réduction de ces inégalités ? Et quels sont leur rôle et leur responsabilité dans la force du lien entre inégalités sociales et inégalités scolaires ?

Ces questions, largement abordées par la sociologie de l'éducation, questionnent le rôle que les systèmes éducatifs sont amenés à jouer dans le traitement des différences entre individus devant l'école et en son sein. S'interroger sur l'équité scolaire c'est donc questionner la justice du

système scolaire, les valeurs et conceptions de la justice qui engagent (ou non) un traitement particulier des inégalités, et donc des différences entre individus jugées comme problématiques au sein d'un système éducatif et d'une société donnés.

Puisque notre recherche s'interroge sur les effets des SABP sur l'équité des systèmes d'éducation, il nous semble essentiel de préciser les figures des inégalités qui seront mobilisées pour pouvoir appréhender cette question. S'intéresser à l'équité scolaire nécessite en effet de se doter d'instruments et de critères à l'aune desquels cette dernière sera mise en débat. Si nous mobilisons trois figures des inégalités – les inégalité des chances de résultats, les inégalités de résultats et les inégalités des acquis de base – pour mettre en perspective les effets des SABP sur l'équité scolaire, il nous semble nécessaire de discuter des implications de ces notions et d'éclairer les principes de justice qui les sous-tendent.

Se tourner vers les travaux de la sociologie de l'éducation, de même que vers les travaux de philosophie politique et les théories de justice qui y sont discutées, s'avère dès lors pertinent pour éclairer ces considérations. Certains travaux de philosophie politique s'intéressent en effet « aux principes normatifs qui devraient régir les institutions de base d'une société » (Pourtois, 2008, p. 49). Ils permettent de penser la justice des systèmes éducatifs au travers de principes structurants. Bien que ces travaux ne traitent pas spécifiquement (pour certains) du champ éducatif, les considérations qui y sont développées offrent un cadre pertinent pour appréhender la question des inégalités et de l'équité scolaires. En effet, pour reprendre les propos de Rawls (2003, p. 25), « le rôle des principes de justice, en tant que composante d'une conception politique de la justice, est de spécifier les termes équitables de la coopération sociale ».

2.7.1 Égalité des chances et méritocratie

Pour Crahay (2012), l'idéal d'égalité des chances, modèle dominant dans la majorité des sociétés occidentales démocratiques, a émergé avec l'école moderne, une fois que l'obligation scolaire eut produit l'essentiel de ses effets, c'est-à-dire la scolarisation de tous les enfants. La notion d'égalité des chances « repose sur l'un des principes normatifs fondamentaux des sociétés modernes : l'égalité de droits entre les Hommes » (Verhoeven & Dupriez, 2008, p. 98). L'idéal d'égalité des chances s'inscrit dans une approche ressourciste de la justice sociale, dans laquelle

s'inscrivent les travaux de Rawls (le ressourcisme rawlsien ²⁶ – (Farvaque, 2008). Pour ce dernier, en effet, une « institution est *juste* lorsqu'elle n'opère aucune distinction *arbitraire* entre personnes dans l'attribution des droits et des devoirs et lorsqu'elle détermine un équilibre *adéquat* entre revendications concurrentes portant sur les avantages de la vie sociale » (Parijs, 1991, p. 71).

« L'approche égalitariste de la justice sociale prend comme principe de base que les diverses ressources doivent être produites et distribuées en accordant le même respect et la même considération aux intérêts de chaque individu (Klymlicka, 1990 ; Swift, 2001). Dans la mesure où ces intérêts diffèrent en raison de la diversité des conceptions de la vie bonne, il importe que chacun puisse disposer de ressources qui lui permettront de réaliser au mieux et par lui-même sa propre conception de la vie bonne. Il en découle une défense du principe d'égalité des chances. La justice exige que l'on égalise les chances dans le choix et la poursuite d'une activité sociale permettant de mener une vie bonne. Relevons que la notion de chance ne désigne pas ici une probabilité statistique, mais l'opportunité, la possibilité, d'atteindre l'objectif que l'on s'est donné » (Pourtois, 2008, p. 51).

Le principe d'égalité des chances repose ainsi sur la légitimité pour chacun d'acquérir une formation de son choix et sur le droit de pouvoir recevoir l'éducation qu'il mérite au-delà de l'origine sociale, du sexe, de la nationalité, de l'origine ethnique ou régionale et des ressources économiques dont il dispose. Chacun doit dès lors pouvoir accéder aux meilleures formations et aux meilleures études, s'il en a la capacité. L'école fonctionne en ces termes selon un principe d'égalité formelle (Duru-Bellat, 2002). Son rôle, ainsi que celui de l'État, est dès lors « d'éliminer les obstacles géographiques, financiers, institutionnels voire psychologiques qui pourraient détourner un enfant ou un adolescent d'une formation qu'il est capable de suivre avec succès » (Perrenoud, 1981, p. 95 cité dans Crahay, 2012, p. 53). Ainsi, selon Grisay :

« une politique éducative inspirée par l'égalité des chances devrait permettre qu'à aptitudes égales, les élèves d'origines modestes aient autant de chances que les enfants de *bonne famille* de fréquenter les filières *nobles* ou d'avoir une scolarité longue » (Grisay, 1984 cité dans Crahay, 2012, p. 55).

²⁶ Certains situent par ailleurs l'idéal d'égalité des chances dans le cadre d'une approche égalisatrice de la justice. Nous ne rentrerons pas dans les nuances qui sont apportées entre les différentes approches égalitaristes/ressources de la justice sociale telles que développées en philosophie politique. Nous renvoyons pour ce faire par exemple à l'ouvrage de Kymlicka (Kymlicka, 1999) qui présente et discute différentes théories de la justice développées en philosophie politique et les approches dans lesquelles elles s'insèrent. Notons par ailleurs que certains auteurs (Voir par exemple Van Parijs, 1991, p. 90-94) émettent des réserves sur la qualification de la théorie de Rawls comme égalitariste. Nous ne trancherons pas cette question ici.

L'idéal d'égalité des chances s'inscrit dans une conception méritocratique de la justice. Cette dernière implique « qu'à l'école comme dans la société chacun soit récompensé ou valorisé selon ses mérites propres » (Crahay, 2013, p. 41). Si la société assure aux individus un accès à la scolarité pour tous les individus au-delà de leur appartenance sociale, la sélection scolaire méritocratique peut alors s'opérer. Ce principe postule et admet la présence d'un « potentiel de formation déterminé » chez les individus (Perrenoud, 1981 cité dans Crahay, 2013, p. 53). La compétition entre individus dans le cadre scolaire doit permettre à chacun d'exprimer ce potentiel et de se positionner au sein de l'échelle scolaire. L'inégalité de résultats est, de fait, admise puisque liée aux capacités et aux potentialités de chacun.

Certaines politiques éducatives structurelles qui ont été mises en place dans divers systèmes d'éducation peuvent ainsi être analysées à la lumière d'une conception méritocratique de l'égalité des chances. Il en va des politiques de démocratisation de l'accès à l'éducation visant la gratuité scolaire et permettant que chaque enfant puisse accéder à une instruction de base, des mesures pour contrer les disparités géographiques et favoriser la construction d'écoles sur un territoire donné et/ou les systèmes de transport pour se rendre à l'école. Ces politiques ou réformes ont été mises en œuvre dans divers contextes afin de favoriser l'égalité d'accès à l'instruction pour tous les élèves. Chacun a donc droit à un accès égal à l'école et à une scolarité de base.

Le principe de l'égalité de traitement peut, par ailleurs, être considéré comme largement inspiré de cet idéal d'égalité des chances (Dupriez & Verhoeven, 2008). Il est en phase avec un idéal d'unification des structures et des parcours scolaires, vers la question de l'école unique pour tous les élèves. Il s'agit d'offrir à tous les élèves les mêmes conditions d'enseignement, entre autres par le biais de structures éducatives unifiées. Il en va, par exemple, des politiques d'homogénéisation institutionnelle des structures éducatives et des parcours scolaires, des contenus d'enseignement aux ordres d'enseignement primaire ou secondaire : *comprehensive schools* en Angleterre, unification de l'école primaire en France en 1959 et du collège unique institué par la loi Haby de 1975, ou encore, des réformes visant l'allongement du tronc commun dans divers systèmes éducatifs.

« Une politique éducative inspirée par l'idéal de l'égalité de traitement vise une égale qualité de l'enseignement dispensé par les écoles relevant d'un même système. Elle s'attache donc nécessairement à homogénéiser les conditions d'enseignement et, donc, à rendre équivalents d'une école à l'autre le taux d'encadrement des élèves, la qualité et la diversité des locaux et des équipements pédagogiques (...) les politiques gèrent de façon égalitaire la qualité des occasions éducatives offertes aux élèves par les établissements scolaires » (Pourtois, 2008, p. 60)

En donnant à tous les élèves la possibilité d'accéder à une même instruction de base et en plaçant les individus dans des conditions égales, ce sont donc les capacités (talents) et l'effort des élèves (le mérite) qui expliqueraient les différences de résultats scolaires. Les différences de résultats des élèves et, donc, les inégalités scolaires n'apparaissent pas comme injustes. Selon une conception méritocratique de la justice, ce qui importe est donc un accès égal aux ressources éducatives et une juste distribution de ces dernières.

Toutefois, la question de l'égalité de traitement va être fortement remise en question par les travaux de la sociologie de l'éducation qui ont mis en exergue différents mécanismes par lesquels l'égalité réelle des chances est en fait altérée au sein de l'institution scolaire. La théorie de la reproduction (Bourdieu, 1966; Bourdieu & Passeron, 1964) démontre en effet que la culture scolaire, teintée de la culture et des codes sociaux des catégories (ou classes) sociales les plus favorisées, tend, de fait à avantager les élèves qui en sont issus. En traitant tous les enseignés sous la même égide, l'institution scolaire serait en fait amenée à reproduire en son sein, et accentuer à sa sortie, les inégalités sociales qui lui préexistent. Cette idée est ainsi au cœur des propos de Bourdieu :

« Pour que soient favorisés les plus favorisés et défavorisés les plus défavorisés, il faut et il suffit que l'école ignore dans le contenu de l'enseignement transmis, dans les méthodes et les techniques de transmission et dans les critères du jugement, les inégalités culturelles entre les enfants des différentes classes sociales : autrement dit, en traitant tous les enseignés aussi inégaux soient-ils en fait, comme égaux en droits et en devoirs, le système scolaire est conduit à donner en fait sa sanction aux inégalités initiales devant la culture » (Bourdieu, 1966, p. 336-337).

La compensation du poids des inégalités d'origines socioéconomiques dans la réussite scolaire, au cœur de la mise en œuvre des politiques compensatoires, est dès lors posée. Il s'agit donc de compenser, par différentes mesures, le poids de différents facteurs et leur influence sur le parcours des élèves.

C'est ainsi qu'États-Unis *Title I* est lancé dans les années 1960, prévoyant l'attribution de financements complémentaires pour la mise en œuvre de mesure de soutien aux élèves d'origine socioéconomique défavorisée. En France, la politique des Zones d'Éducation prioritaires s'inscrit aussi dans cette logique compensatoire (Rochex & Kherroubi, 2004). En outre, la mise en œuvre de politiques compensatoires ne peut être déconnectée de la définition de l'échec scolaire comme problème social. Ainsi, comme l'a pointé Hutmacher, l'emploi de la notion d'échec scolaire « suppose implicitement que l'on s'attend à ce que tous réussissent ou puissent réussir » (Hutmacher, 1993, p. 31). Ce construit social implique des attentes différentes envers l'école et son rôle dans la réussite des élèves. :

« Il n'y a pas d'échec tant que l'on ne s'attend pas à une réussite, tant que l'explication renvoie l'incapacité disciplinaire ou intellectuelle de certains élèves à des caractéristiques individuelles sur lesquelles l'école ne se donne pas prise » (Hutmacher, 1993, p. 30)

2.7.2 Égalité de résultats et égalité des acquis de base

Le concept d'égalité de résultats vient opérer un déplacement dans le principe de justice sous-jacent. Dupriez et Verhoeven parlent ici de la traduction, dans cet objectif d'égalité de résultats, d'un point de vue égalitariste strict de la justice « considérant qu'une société juste c'est une société sans inégalités » (Dupriez & Verhoeven, p. 109, 2008).

Ce qui est visé n'est plus seulement une égalité ou une égalisation des ressources éducatives et des conditions d'enseignement, mais des acquisitions effectives des élèves. Il ne s'agit donc plus uniquement de donner aux élèves les mêmes chances de réussite, mais de faire en sorte que « tous » les élèves réussissent. L'égalité de résultats implique donc de s'intéresser à la distribution des résultats à l'issue d'une période de scolarité et de formation données. L'école doit alors jouer un rôle et se fixer des objectifs plus ambitieux dans le traitement des inégalités :

« L'école doit être à même de compenser les inégalités d'aptitudes elles-mêmes (quelque soit leur origine, génétique ou familiale) (...) on glisse ainsi du concept d'égalité des chances ou de l'égalité d'accès à l'égalité de résultats, faisant le pari que l'école peut et doit avoir ce rôle actif, radical pour créer l'égalité de réussite entre enfants » (Duru-Bellat, 2002, p. 13).

La notion d'égalité des acquis de base²⁷ reprend cette idée d'égalisation des résultats à l'issue d'une période de scolarité donnée. Toutefois, ce qui est visé n'est pas les résultats égaux de l'ensemble des élèves, mais la fixation d'un seuil minimal que chaque individu doit pouvoir atteindre. L'égalité des acquis de base prône ainsi un « traitement contrôlé et orienté des différences individuelles vers l'acquisition d'un niveau commun de base » (Dupriez & Verhoeven, 2008, p. 60). La fixation de ce seuil engage alors diverses considérations : s'agit-il d'un seuil de connaissances et/ou compétences? S'agit-il d'un niveau de formation donné? Comment s'opère la fixation de ce seuil au sein de chaque société ou de chaque système éducatif et quels critères entrent alors en ligne de compte ?

La notion d'égalité des acquis de base entraîne donc la nécessaire définition de ces acquis de base. Ils peuvent être considérés comme un ensemble de savoirs et de compétences que les individus doivent posséder au sein d'une société donnée afin d'être en mesure d'y participer et d'y contribuer. En somme, il faut outiller l'ensemble des élèves d'un bagage commun au sortir de l'école.

La notion d'égalité des acquis de base implique de fait une vision moins ambitieuse que celle d'égalité de résultats. Cette dernière correspond « au sens strict » « à une situation où tous les élèves obtiennent des performances égales à l'issue d'une période de formation » (Dupriez & Vandenberghe, 2008, p. 168).

La discussion de ces différentes figures de l'égalité (des chances, de résultats, des acquis de base) et les conceptions de justice qui les sous-tendent permettent donc de mettre en exergue le fait que ces dernières engagent des conceptions relativement différentes des inégalités « acceptables » au sein d'une société donnée. Ainsi, la notion d'inégalité peut, selon la définition qu'on en donne, renvoyer à des principes de justice fort différents²⁸.

²⁷ Pour Dupriez et Verhoeven, la notion d'égalité des acquis de base, et plus particulièrement la fixation d'un socle commun de compétence en Belgique francophone, renvoie en partie à une approche « sufficientiste » de la justice qui considère que « tous les individus doivent pouvoir bénéficier des ressources qui leur permettent d'assurer un certain nombre de besoins vitaux. Et l'éducation mérite d'être considérée comme une ressource indispensable à ces besoins » (Dupriez & Verhoeven, 2008, p. 110).

²⁸ Il est aussi nécessaire de préciser que les conceptions de l'égalité ne sont pas mutuellement exclusives au sein d'un système éducatif. Elles peuvent différer en fonction des ordres d'enseignement (primaire, secondaire, universitaire) : on peut ainsi revendiquer l'égalité des acquis pour l'enseignement primaire et valoriser davantage une justice méritocratique aux ordres d'enseignement secondaire supérieur et universitaire.

Ces conceptions de la justice ont été et sont mobilisées de manières variables dans les systèmes éducatifs, en fonction du contexte historique, social et politique car, pour reprendre les propos de Duru-Bellat (2002), la définition d'une inégalité comme un problème est toujours liée à un contexte sociopolitique particulier. Ces conceptions de justice peuvent donc être reprises dans les politiques éducatives. Elles alimentent « des propositions, mais aussi des critiques et des dénonciations de l'action en éducation et en formation » (Dupriez, Oriane & Verhoeven, 2008, p. 161).

Au total, si la mise en œuvre des SABP s'appuie sur une rhétorique de l'équité, leurs potentialités à altérer cette dimension seront analysées dans notre thèse doctorale en convoquant trois figures de l'égalité (et des inégalités) qui, nous l'avons vu, renvoient finalement à des principes de justice différents : 1) les inégalités des chances de résultats, 2) les inégalités de résultats et 3) les inégalités des acquis de base.

1) La notion d'inégalité des chances de résultats renvoie ainsi à la force du lien entre les résultats scolaires et l'origine socioéconomique des élèves au sein des systèmes éducatifs. En d'autres termes, elle pose la question des chances de réussite des élèves en fonction de leur milieu socioéconomique. Elle vise donc à interroger la participation des SABP à l'idéal d'égalité des chances discuté dans cette sous-section.

2) L'égalité de résultats réfère à l'atteinte pour les élèves d'un même niveau de résultats. Son corollaire, l'inégalité de résultats à la non-atteinte d'un même niveau de résultats. Elle permet de discuter à la capacité des SABP à tendre vers cet objectif. Nous nous intéresserons ici au différentiel de résultats entre élèves, plus particulièrement entre les élèves les plus performants et les élèves les plus faibles académiquement.

3) Enfin, l'inégalité des acquis de base est entendue comme l'inégale accession à un niveau de compétences de base. Il nous semble d'autant plus pertinent de mobiliser cette notion que l'atteinte d'un seuil minimum de connaissances et de compétences est un objectif fortement valorisé dans les systèmes d'éducation au cours des dernières décennies. De plus, la fixation d'un seuil minimum de connaissances pour tous les élèves et la mesure de l'atteinte de cet objectif ont été au cœur de la mise en œuvre des SABP, du moins dans certains contextes.

2.8 Conclusion

Cette section nous a permis de situer conceptuellement les notions d'efficacité et d'équité. Nous avons donc mis en exergue les enjeux qui entourent ces notions dans le champ éducatif et avons par ailleurs montré que la notion d'équité scolaire renvoie finalement à la nécessaire discussion de différentes figures des inégalités, et donc à différentes conceptions de la justice sociale et scolaire.

Nous avons par ailleurs arrêté les acceptions de ces concepts centraux au sein de notre recherche. Ainsi, nous interrogeons les effets des SABP d'une part sur le degré d'efficacité scolaire des systèmes éducatifs qui renvoie, dans notre travail doctoral, aux performances moyennes des élèves dans les systèmes d'éducation. Les effets des SABP sur l'équité scolaire seront aussi analysés en nous penchant sur trois types d'inégalités : l'inégalité des chances de résultats, l'inégalité de résultats et l'inégalité des acquis de base. Le choix de mobiliser ces trois types d'inégalités repose sur le constat dressé précédemment d'une conception relativement restreinte de l'équité scolaire mobilisée dans les recherches s'intéressant aux effets des SABP dans le champ éducatif.

Au total, en mobilisant les concepts d'efficacité et d'équité scolaires analysées de manière concomitante, nous nous rangeons de fait derrière nombre d'auteurs qui pointent la nécessité de questionner les potentialités des politiques éducatives à la lumière de ces deux notions, dès lors convoquées de concert (Mons, 2017).

2.9 Synthèse et démarche générale de recherche

Ce chapitre nous a permis de situer conceptuellement les systèmes d'*accountability* basés sur la performance au travers des concepts de régulation institutionnelle et de modes de régulation institutionnelle. Nous avons pu éclairer les définitions de ces notions, mobiliser la littérature sur les instruments d'action publique pour mettre en exergue l'acception de l'*accountability* et celle des SABP au cœur de notre travail de recherche doctorale. Nous avons par ailleurs soutenu la pertinence d'une entrée d'analyse par les instruments d'action publique pour caractériser les SABP et analyser leurs effets sur l'efficacité et l'équité scolaires.

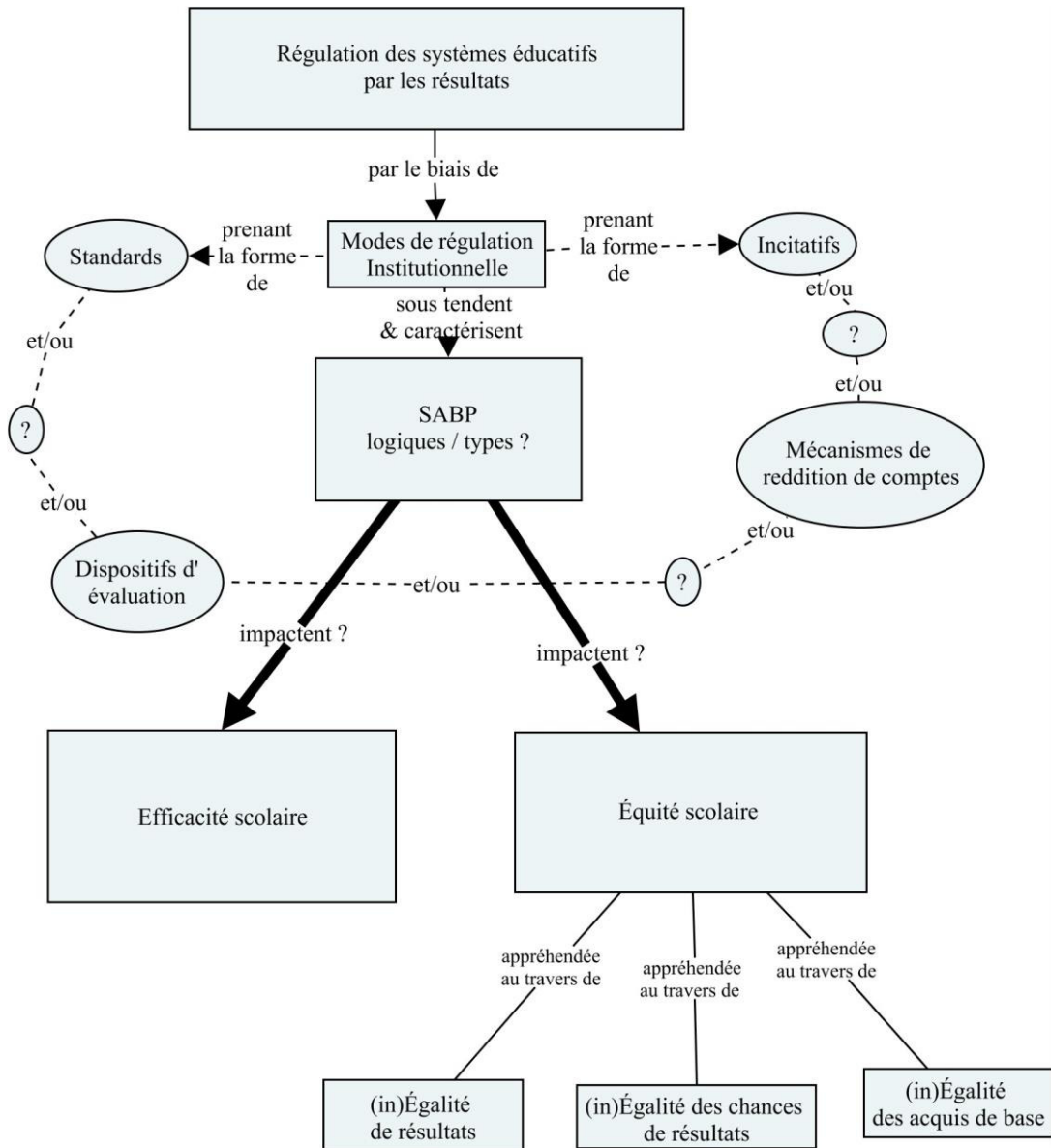
La deuxième section de ce chapitre a permis de stabiliser les notions clefs d'efficacité et d'équité scolaires. Nous y avons explicité les enjeux qui sont attachés à ces notions dans le champ

éducatif. De même, nous y avons soulevé la nécessité d'un effort de conceptualisation où différentes conceptions de l'égalité seraient prises en compte pour problématiser et définir la notion d'équité scolaire.

De fait, nous avons éclairé les principes de justice qui sous-tendent différentes figures de l'égalité, et donc, la définition des inégalités scolaires considérées comme inacceptables au sein des systèmes d'éducation. La notion d'équité scolaire implique en effet un jugement normatif sur le caractère juste ou injuste des différences entre individus au sein de l'institution scolaire. Nous avons ainsi précisé quelles sont les figures de l'égalité rapportées aux résultats des élèves interrogées dans notre travail de recherche : l'égalité de résultats, l'égalité des chances de résultats et l'égalité des acquis de base.

Notre recherche vise donc à analyser le lien entre SABP, efficacité et équité scolaires. Elle pose, en effet, la question des effets des SABP sur le degré d'efficacité et d'équité scolaire. La Figure 3 permet d'illustrer le schéma conceptuel au cœur de notre recherche.

Figure 3 : Schéma conceptuel systèmes d'*accountability* basés sur la performance (SABP), efficacité et équité scolaires



Les principaux concepts de la recherche étant posés, les relations entre ces derniers illustrées, il nous est dès lors possible d'exposer les perspectives générales et spécifiques de la recherche. Ces perspectives sont établies en lien avec les éléments de problématisation évoqués au chapitre précédent, les concepts au cœur de notre recherche et le questionnement central de notre thèse doctorale.

Nous avons mis précédemment en exergue un manque dans la littérature : des études quantitatives comparatives inter-systèmes d'éducation qui permettent d'analyser les effets des SABP sur l'efficacité et l'équité scolaires, en dehors du contexte états-unien, sont peu nombreuses et présentent des lacunes.

Nous avons par ailleurs souligné le besoin d'un effort de conceptualisation et d'opérationnalisation des SABP, compte tenu du faible nombre d'études internationales s'intéressant à cette question, et de la nécessité de dépasser la distinction SABP à forts et à faibles enjeux sur laquelle la littérature comparative empirique s'est majoritairement appuyée. De même, cet effort est aussi nécessaire en ce qui a trait à la conceptualisation et à l'opérationnalisation de la notion d'équité scolaire.

Finalement, à l'instar de plusieurs auteurs, nous soutenons que l'efficacité et l'équité scolaires doivent être analysées de concert, particulièrement lorsqu'il s'agit d'étudier les effets et les potentialités des politiques éducatives (Attali & Bressoux, 2002; Mons, 2017), et plus spécifiquement, des SABP dans le champ éducatif.

À la lumière de ces éléments, la question de recherche centrale au cœur de notre recherche doctorale peut-être formulée de la sorte : quels sont les effets des systèmes d'*accountability* basés sur la performance sur les degrés d'efficacité et d'équité des systèmes éducatifs ?

Elle appelle les sous-questions de recherche suivantes : 1) quelles formes revêtent les SABP dans les systèmes éducatifs d'Europe et du Canada ? 2) Les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ?

Ces deux sous-questions feront l'objet de deux volets empiriques de recherche poursuivant des objectifs distincts, mais interreliés. Leur réalisation nous permettra ainsi de produire des évidences empiriques permettant d'apporter des éléments de réponse à notre questionnement central.

Le premier volet de recherche poursuit tout d'abord un objectif descriptif et analytique. Il s'agit d'une part 1) de caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada, d'autre part 2) de classer les systèmes éducatifs de notre échantillon sur cette base. En termes opérationnels, il vise donc la création de variables originales permettant de caractériser les SABP dans les systèmes éducatifs de notre échantillon et d'en mesurer les effets sur l'efficacité et l'équité scolaires (dans un deuxième volet de recherche).

Il mobilise plus particulièrement le concept de SABP qui y sera opérationnalisé en cohérence avec notre cadrage théorique et l'entrée par l'instrumentation que nous privilégions.

Plusieurs postulats en lien avec notre cadrage théorique guident ainsi la réalisation de ce premier volet empirique. Nous postulons que les instruments au cœur des SABP sous-tendent des logiques instrumentales de régulation distinctes et nous permettent de différencier par ailleurs plusieurs types empiriques de SABP au sein desquels il est possible de regrouper les systèmes éducatifs de notre échantillon.

Le second volet de recherche vise quant à lui à analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires. L'hypothèse au cœur de ce volet de recherche est celle de l'effet différencié des SABP sur l'efficacité et l'équité scolaires en fonction des logiques privilégiées et des types empiriques distincts mis en exergue dans le premier volet de recherche.

Les concepts de SABP, d'efficacité et d'équité scolaires seront plus particulièrement au cœur de ce volet qui nous permettra d'analyser la relation et les effets des SABP sur les degrés d'efficacité des systèmes éducatifs de notre échantillon.

Le tableau synoptique du design de recherche est placé ci-dessous (Tableau 1). Nous y reprenons la question centrale de la thèse, les sous-questions et les objectifs de recherche. Les postulats et les hypothèses de recherche sont aussi mentionnés, de même que les concepts clefs sur lesquels s'appuient notre recherche. Finalement, nous y reprenons le type de recherche, les sources de données, l'échantillon et les méthodes d'analyse de données, éléments que nous discutons dans le chapitre suivant dédié à la méthodologie de la recherche.

Tableau 1 : Tableau synoptique du design de recherche

	Volet empirique 1	Volet empirique 2
Question centrale de la thèse	Quels sont les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation ?	
Sous-questions de recherche	Quelles formes revêtent les SABP dans les systèmes d'éducation d'Europe et du Canada ?	Les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ?
Objectifs	<u>Descriptifs et analytiques</u> : 1) caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada ; 2) classifier les systèmes d'éducation de l'échantillon sur cette base <u>Opérationnel</u> : 2) construction de variables permettant d'analyser les effets des SABP sur l'efficacité et l'équité scolaires	<u>Analytique</u> : 1) analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires
Postulats/hypothèses	1) Les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes. 2) Les instruments permettent de différencier des types empiriques de SABP.	2) Effets différenciés des SABP (composantes, types empiriques) sur l'efficacité et l'équité scolaires
Concepts clefs	Modes de régulation institutionnelle, SABP	SABP, efficacité scolaire, équité scolaire (inégalités de résultats, inégalités des chances de résultats et inégalités des acquis de base)
Type de recherche	Étude comparative quantitative internationale	
Données	PISA 2012	
Échantillon	Systèmes éducatifs d'Europe et du Canada	
Méthodes d'analyse des données	Analyses multivariées	Analyses bivariées et multivariées

CHAPITRE 3 : METHODOLOGIE

Ce chapitre présente les éléments de méthodologie au cœur de la recherche. Il permet de décrire le dispositif empirique déployé, l'opérationnalisation des concepts centraux discutés dans le chapitre précédent : les SABP, l'efficacité et l'équité scolaires (analysée sous l'angle de trois figures des inégalités : l'inégalité de résultats, l'inégalité des chances de résultats, l'inégalité des acquis de base).

Dans ce chapitre nous situons tout d'abord notre recherche dans les cadres de l'éducation comparée avant de présenter les données mobilisées (principalement celles du PISA 2012), ainsi que l'échantillon sur lequel porte notre recherche. Dans les sous-sections suivantes nous présentons la démarche méthodologique et les méthodes d'analyse des données. Puisque nous déployons deux volets empiriques de recherche, le premier visant à caractériser les SABP dans les systèmes éducatifs de notre échantillon, le second à analyser leurs effets sur l'efficacité et l'équité scolaires, nous en précisons les éléments essentiels (opérationnalisation des concepts centraux, variables mobilisées, méthodes d'analyses des données).

3.1 Une étude comparative quantitative internationale

Selon Van Daele (1993), l'éducation comparée peut être définie en ces termes :

« a) la composante pluridisciplinaire des sciences de l'éducation ; b) qui étudie des phénomènes et des faits éducatifs ; c) dans leur relation avec le contexte social, politique, économique, culturel, etc. ; d) en comparant leurs similitudes et leurs différences dans deux ou plusieurs régions, pays, continents ou au niveau mondial ; e) afin de mieux comprendre le caractère unique de chaque phénomène dans son propre système éducatif et de trouver des généralisations valables ou souhaitables ; f) dans le but final d'améliorer l'éducation » (Van Daele, 1993, p. 16).

Situer notre recherche dans le champ de l'éducation comparée s'impose dès lors comme une évidence. La définition que propose Van Daele (1993) peut en effet être aisément rapportée aux objectifs de notre recherche. Nous souhaitons analyser les SABP dans divers systèmes d'éducation et mettre en exergue leurs similitudes et leurs différences en ce qui a trait aux instruments sur lesquels ces derniers reposent.

Par ailleurs, puisque nous souhaitons interroger les effets des SABP sur les degrés d'efficacité et d'équité scolaires, la pertinence du choix d'une posture comparatiste apparaît clairement dans les propos de Durkheim :

« Nous n'avons qu'un seul moyen de démontrer qu'un phénomène est cause d'un autre, c'est de montrer les cas où ils sont simultanément présents ou absents et de chercher si les variations qu'ils présentent dans ces différentes combinaisons de circonstances témoignent que l'un dépend de l'autre. Quand ils peuvent être artificiellement produits au gré de l'observateur, la méthode est l'expérimentation proprement dite. Quand, au contraire, la production des faits n'est pas à notre disposition et que nous ne pouvons que les rapprocher tels qu'ils se sont spontanément produits, la méthode que l'on emploie est celle de l'expérimentation indirecte ou méthode comparative » (Durkheim, 1894, p. 168).

Selon Groux et Porcher, l'approche comparative permet « d'approcher les réalités éducatives internationales, de les étudier par le biais de la comparaison, dans leur contexte global » (Groux et Porcher, 1997, p. 117 cités dans Mons, 2004). De fait, cette approche est d'autant plus pertinente que le développement des SABP dans le champ éducatif n'est pas un phénomène isolé ou une spécificité nationale, mais bien une réalité qui touche la plupart des systèmes d'éducation des pays du globe. La posture comparative inscrit d'une part de fait notre objet dans les tendances globales qui traversent le champ éducatif, mais permet aussi et d'autre part d'apprécier les variations des SABP (leurs outils, leurs formes différenciées) en fonction des particularités de contexte.

Mons le précise, la posture comparatiste permet en effet de saisir la « variété des modes d'organisation des systèmes d'éducation » (2017, p. 307), difficilement observables dans les contextes nationaux. Les études quantitatives comparatives et la sociologie quantitative internationale l'ont par ailleurs bien montré, la variation de certains facteurs clefs (économiques, structurels et organisationnels) permet en partie d'expliquer les différences en termes d'efficacité et d'équité scolaires des systèmes d'éducation (Marks, 2005; Mons, 2004, 2017). Ainsi, en éclairant l'influence « relative » des facteurs scolaires sur ces dimensions, elles ont mis en exergue « les potentialités des politiques éducatives » (Mons, 2004, p. 38). Finalement, les études comparatives internationales revêtent une valeur heuristique non négligeable :

« Grâce à la diversité des situations éducatives qu'elles apportent (dans le cadre d'études à couverture internationale plutôt que régionale), elles ont permis de mettre en évidence des relations qui n'avaient pas été décelées jusqu'alors dans des études comparatives intra-systèmes » (Mons, 2004, p. 39)

Précisons que nous ne mobilisons pas la comparaison pour saisir le caractère original ou la spécificité culturelle des systèmes d'éducation comme le voudrait le paradigme particulariste (Mons, 2004; Raveaud, 2007). Si nous ne nions pas pour autant les spécificités de chaque système éducatif, notre recherche vise davantage à interroger la relation entre SABP, efficacité et équité scolaires. Notre recherche se rapproche donc davantage des études situées dans le paradigme universaliste, elle peut-être qualifiée de « *cross national* » (Hantrais, 1999; Maurice, 1989) et vise à identifier des régularités, des relations entre variables spécifiées. Nous rejoignons ainsi la définition de l'éducation comparée que proposent Noah et Eckstein :

« L'éducation comparée est une tentative d'utiliser des données internationales pour tester des propositions sur les relations qui peuvent exister entre éducation et société ainsi qu'entre pratiques pédagogiques et résultats scolaires » (Noah & Eckstein, 1969 cités dans Mons, 2004, p. 34).

Cependant, au travers de notre démarche, nous ne prétendons pas à l'établissement de lois générales. En cela, nous nous rangeons derrière ceux qui postulent que la diversité des contextes et des situations dans le champ éducatif rend périlleux l'établissement de lois générales.

À l'instar de Holmes (1981), nous pensons qu'il est préférable de traiter (au mieux) de « généralisations contextualisées dans lesquelles les comportements sont gouvernés par des conditions spatiotemporelles, sociopsychologiques, linguistiques ou économiques spécifiques » (Cité dans Mons, 2004, p. 56).

L'importance des spécificités nationales est, en outre, l'un des éléments qui soutiennent la pertinence d'une analyse comparative des effets des SABP qui dépasse le cadre états-unien où prennent place la plupart des études sur la question. Il s'agit d'élargir le spectre de l'analyse, d'observer les formes différenciées des SABP dans d'autres contextes, d'analyser dans quelle mesure ces formes différenciées impactent, le cas échéant, les degrés d'efficacité et d'équité scolaires.

Puisque les SABP revêtent des types différenciés au sein des systèmes d'éducation d'un même pays, comme en témoignent les études dans le contexte états-unien, notre unité d'analyse est le système éducatif. Nous avons donc souhaité, lorsque cela était permis par nos données, prendre en compte ces variations au sein d'un même pays.

Finalement, la comparaison synchronique (dans l'espace plutôt que dans le temps) termes à

termes nous permet de comparer les instruments qui sous-tendent et mettent en acte les SABP. En d'autres termes, nous nous intéressons aux configurations d'instruments à un instant donné dans les systèmes éducatifs de notre échantillon. Nous souhaitons observer si les variations inter-systèmes éducatifs, en termes de modes de régulation mis en œuvre au sein des établissements scolaires, influencent de manière différenciée leurs degrés d'efficacité et d'équité scolaires.

Le cadre méthodologique général étant précisé, nous présentons, dans la sous-section suivante, les sources de données mobilisées dans notre étude.

3.2 Les données

« Il faut bien connaître les principes qui ont présidé à la confection [des statistiques]. Faute de précautions minutieuses, on risque d'aboutir à des données fausses : ainsi il est impossible d'utiliser les renseignements statistiques sur le suicide en Angleterre, car, dans ce pays, pour éviter les rigueurs de la loi, la plupart des suicides sont déclarés sous le nom de mort par suite de folie ; la statistique est ainsi viciée dans son fondement » (Mauss, 1968, p. 33).

Notre recherche repose principalement sur une analyse secondaire des données du *Programme for International Student Assessment* (PISA) 2012 pilotée par l'Organisation de Développement et de Développement Économique (OCDE). Nous avons aussi consulté les bases de données internationales pour compléter certaines informations non disponibles dans cette base de données et nous nous sommes référée à la documentation internationale et nationale pour obtenir des renseignements supplémentaires pertinents.

3.2.1 L'enquête PISA 2012

Les données du PISA 2012 constituent notre source principale d'information. Cette enquête internationale cyclique, dont la première édition a été lancée en 2000, est pilotée par l'Organisation de Développement et de Développement Économique (OCDE). Elle a pour vocation de « déterminer dans quelle mesure les élèves ont acquis, au terme de leur scolarité obligatoire, les savoirs et les savoir-faire nécessaires à leur pleine participation à la “société de la connaissance” » (Le Donné, 2017, p. 17). Ce programme, qui vise à évaluer les compétences des jeunes scolarisés de 15 ans, a lieu tous les trois ans. Il est décrit ainsi par l'OCDE (OCDE, 2014c):

« Qu'importe-t-il de savoir et de savoir-faire en tant que citoyen ? » C'est la question qui sous-tend le système mondial d'évaluation de la qualité, de l'équité et de l'efficacité de l'éducation, en l'occurrence le Programme international pour le suivi des acquis des élèves (PISA). L'enquête PISA évalue dans quelle mesure les élèves âgés de 15 ans ont acquis des connaissances et compétences essentielles pour pouvoir participer pleinement à la vie de nos sociétés modernes, en particulier en compréhension de l'écrit, en mathématiques, en sciences et en résolution de problèmes. L'enquête PISA ne cherche pas simplement à évaluer la faculté des élèves à reproduire ce qu'ils ont appris, mais vise aussi à déterminer dans quelle mesure ils sont capables de se livrer à des extrapolations à partir de ce qu'ils ont appris et d'utiliser leurs connaissances dans des situations qui ne leur sont pas familières, qu'elles soient ou non en rapport avec l'école. Cette approche reflète le fait que les sociétés modernes apprécient les individus moins pour leurs connaissances que pour leur capacité à utiliser ces dernières » (OCDE, 2014, p. 3)

Le programme PISA s'appuie sur un ensemble de principes. Premièrement, il revêt une orientation politique, puisque « le cadre conceptuel et les méthodes de compte rendu de l'évaluation ont été choisis pour permettre aux gouvernements de tirer des enseignements politiques des résultats » (OCDE, 2001, p. 18). Il se focalise par ailleurs sur une manifestation des connaissances et des compétences considérées comme clefs pour évoluer au sein des sociétés modernes et « dans la vie de tous les jours ». Il n'est donc pas aligné sur les curriculums nationaux, mais porte plutôt sur un ensemble de connaissances et de compétences définies par l'OCDE comme centrales dans des sociétés inscrites dans une économie mondialisée. Il repose finalement sur un principe de collaboration puisqu'il est dirigé conjointement par les pays et les économies participants sous l'égide « de l'expertise scientifique d'un consortium international constitué des plus éminents instituts spécialisés dans l'évaluation » (OCDE, 2001, p. 18).

Le programme PISA a connu un succès retentissant, comme en témoigne l'augmentation du nombre de pays participants entre la première et la dernière édition : 30 pays pour la première édition PISA 2000 contre 72 pays et économies participants pour l'édition 2015 (OCDE, 2001, 2016). Chaque édition évalue trois disciplines — la compréhension de l'écrit (ou littérature), la culture mathématique et les sciences. À chaque édition, le focus est porté sur une discipline, « domaine majeur » de l'enquête : compréhension de l'écrit en 2000, culture mathématique en 2003, sciences en 2006, compréhension de l'écrit en 2009.

Malgré les débats et les critiques qui entourent le programme, les enquêtes PISA constituent l'une des sources de données mondiales les plus conséquentes et les plus complètes en éducation. Elles centralisent des informations sur les acquis des élèves, recueillent un ensemble de données

contextuelles « sans précédent » sur les élèves, les établissements et les systèmes scolaires (Le Donné, 2017, p. 21). Les données sont, en outre, publiques et accessibles sur le portail de l'OCDE.

Notre recherche doctorale s'appuie sur les données de l'édition 2012. Le domaine majeur est ici la culture mathématique. Ainsi, 34 pays membres de l'OCDE (soit la totalité des pays membres) et 31 économies partenaires ont participé à cette édition. Environ 510 000 élèves « représentatifs des quelque 28 millions d'élèves âgés de 15 ans²⁹ scolarisés dans les 65 pays et économies participants » (OCDE, 2014b, p. 3) ont passé les tests du PISA. Les critères de sélection des élèves vont comme suit :

“The international PISA target population in each participating country and economy consisted of 15-year-old students attending educational institutions in grade 7 and higher. This meant that countries were to include 15-year-old students: - enrolled full-time in educational institutions; - enrolled in educational institutions who attended only on a part-time basis; - enrolled in vocational training programmes, or any other related type of educational programmes; - and, attending foreign schools within the country (as well as students from other countries attending any of the programmes in the first three categories). It was recognised that no testing of 15-year-olds schooled full-time in the home, workplace or out of the country would occur and therefore these 15-year-olds were not included in the international target population. The operational definition of an age population directly depends on the testing dates. The international requirement was that the assessment had to be conducted during a 42-day period, referred to as the testing period, between 1 March 2012 and 31 August 2012, unless otherwise agreed” (OCDE, 2014a, p. 66)

L'échantillon des pays et des économies participants est constitué sur une base volontaire, celui des écoles et des élèves sur une base probabiliste aléatoire. Le PISA propose un échantillonnage par strate (niveau école, niveau élèves), les critères d'inclusion et d'exclusion sont établis pour chaque strate de l'échantillon (niveau établissements et élèves)³⁰. Il requiert un minimum de 150 écoles participantes dans chaque pays et dans chaque économie. Un minimum de 4 500 élèves par pays doit être inclus à l'échantillon. Des taux de réponse de 50 à 80 % dans les écoles sélectionnées sont requis. L'enquête PISA détermine par ailleurs des seuils critiques en termes de

²⁹ La fourchette va de 15 ans et 3 mois à 16 ans et 2 mois pour cette édition (OCDE, 2014).

³⁰ Nous présentons ici les caractéristiques générales des données du PISA pour introduire celles que nous utilisons. Pour de plus amples informations sur cette enquête (constitution de l'échantillon — procédures, critères d'inclusion et d'exclusion, etc. —, construction des questions, modalités de passation des questionnaires, traductions et limitation des biais culturels, etc.) nous renvoyons au rapport technique du PISA 2012 (OCDE, 2014) disponible à <https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf>

taux de participation et de réponses à partir desquels les biais d'échantillonnage sont mesurés et évalués (OCDE, 2014a).

En plus des épreuves visant à évaluer les acquis des élèves³¹, des questionnaires contextuels, à destination des élèves et des chefs d'établissements, permettent de recueillir des informations supplémentaires sur les élèves, sur leur « milieu familial, leur établissement d'enseignement et leurs expériences concernant leur établissement et leurs apprentissages » (OCDE, 2016, p. 3). Au niveau des établissements scolaires, le questionnaire à l'attention des chefs d'établissements vise à recueillir un ensemble d'informations sur les écoles — climat de l'école, situation de concurrence avec d'autres écoles, sources de financement, ressources, pratiques éducatives et de gestion — et sur les systèmes d'éducation — répartition des compétences entre le niveau central, intermédiaire, local et degrés d'autonomie des écoles, part du financement public et privé, stratification du système (orientation, pratiques du redoublement, etc.).

Nous mobilisons pour nos analyses les données issues des tests des élèves, des questionnaires contextuels « élèves » et « écoles », qui renseignent sur des informations essentielles pour notre recherche. Au niveau des élèves, il s'agit d'abord de leurs performances dans les différentes disciplines testées, ainsi que des informations relatives, par exemple, à leur milieu socioéconomique. Au niveau école, il s'agit principalement d'informations ayant trait aux instruments centraux des SABP (publication des résultats des élèves, suivi des performances des écoles, etc.). Nous mobilisons aussi cette base de données pour nous renseigner sur d'autres facteurs au niveau du système d'éducation — pratique du redoublement, variation du niveau socioéconomique des élèves (etc).

Puisque nous travaillons au niveau du système d'éducation et que les informations sont collectées par PISA au niveau des établissements et des élèves, nous réalisons des procédures permettant d'agrèger les données au niveau du système éducatif. Cette démarche d'agrégation,

³¹ Les évaluations des élèves prennent la forme d'un questionnaire papier crayon pour les mathématiques, l'écrit, les sciences et la « littératie financière », option prévue à l'édition 2012. Comme évoqué précédemment, les tests portent sur un ensemble de connaissances et de compétences, entendues comme la capacité à mobiliser les connaissances dans un ensemble de situations proposées aux tests. Les tests comprennent des questions à choix multiples et ouvertes. La durée du test est d'environ deux heures. Quarante minutes sont prévues pour les épreuves informatisées optionnelles. Cette évaluation assistée par ordinateur porte sur la résolution de problème, les mathématiques et la compréhension de l'écrit. La synthèse des domaines d'évaluation dans les trois matières pour l'édition 2012 est placée en annexe (ANNEXE 3).

que nous expliciterons plus en détails, n'est pas sans poser une question centrale : dans quelle mesure les pratiques déclarées par les établissements et les renseignements collectés à ce niveau d'action peuvent-ils être considérés comme « valides » pour témoigner de la réalité d'un système éducatif ?

Rappelons ici que cette inférence est au cœur de la démarche de l'OCDE ainsi que de la majorité des recherches qui mobilisent les données du PISA pour étudier l'effet de facteurs structurels et organisationnels sur les performances des élèves (Felouzis, 2009; Felouzis, Charmillot, & Fouquet-Chauprade, 2009, 2011; Marks, 2005; Mons, 2004; Rocher & Le Donnée, 2012). Dans notre cas, les données récoltées au niveau des établissements scolaires présentent un intérêt certain. Elles renseignent sur les usages par les établissements d'un ensemble d'instruments et de dispositifs clefs pour notre propos. Elles nous permettent ainsi d'analyser les instruments en action (Levacic, 2004 cité dans Dupriez & Dumay, 2011), de caractériser les SABP au niveau des systèmes éducatifs sur cette base et de questionner leurs effets sur les degrés d'efficacité et d'équité scolaires. En effet, les données du PISA ont la vocation de lier les résultats des élèves à différentes variables structurelles, contextuelles et organisationnelles, objectif central poursuivi par l'OCDE et au cœur de notre démarche d'analyse.

3.2.2 Les bases de données internationales, la littérature institutionnelle et empirique

Nous avons aussi mobilisé les bases de données internationales pour obtenir des informations qui n'étaient pas fournies par le PISA 2012. Nous avons donc consulté les bases de données de la Banque Mondiale ou celles d'EUROSTAT pour construire certaines variables mobilisées dans nos analyses (voir infra).

En complément, nous nous sommes référée à la documentation produite par l'OCDE (OCDE, 2013a, 2013b, 2014c), par Statistique Canada pour les provinces canadiennes (Council, 2013), par l'institut allemand chargé de la production de statistiques en éducation au niveau fédéral (Statistisches-Bundesamt, 2011, 2012).

Enfin, nous nous sommes aussi référée à la littérature institutionnelle internationale (documentation d'Eurydice de la Commission européenne : (Eurydice, 2008, 2009, 2015)) et nationale, ainsi qu'à la littérature empirique sur les SABP dans les systèmes d'éducation intégrés

à notre échantillon pour étayer la description de certains résultats. Le tableau ci-dessous reprend les principales sources de données mobilisées.

Tableau 2 : Sources de données mobilisées

Bases de données	Littérature institutionnelle internationale	Littérature institutionnelle nationale	Littérature scientifique
PISA 2012	OCDE	Conseil des ministres en éducation du Canada	Études empiriques
Banque Mondiale	Statistique Canada	Ministères de l'Éducation des pays (le cas échéant, des provinces ou juridictions)	
EUROSTAT	EURYDICE		
	Commission Européenne		

3.3 L'échantillon

Pour constituer notre échantillon, nous avons adopté une stratégie de sélection qui se rapproche du *Most Different System Design* (Faure, 1994). Cette stratégie vise à maximiser le nombre de cas pour établir des relations statistiques entre variables d'intérêt (Faure, 1994). La variation inter-cas est donc recherchée sur les facteurs centraux à l'étude — dans notre cas, les SABP et leurs instruments.

Nous avons précédemment déploré le fait que peu d'études quantitatives internationales se sont intéressées aux effets des SABP hors contexte états-unien et que peu d'entre elles ont réalisé un effort de typification empirique. Nous avons donc souhaité élargir le spectre de l'analyse aux systèmes éducatifs des pays d'Europe et du Canada.

Si nous avons souhaité inclure un nombre important de systèmes d'éducation à notre échantillon, nous avons cependant veillé à ce qu'ils s'inscrivent dans des « réalités » sensiblement similaires : niveau de développement économique, structures scolaires développées et stables. Nous avons ainsi écarté les pays dont le niveau de développement économique et scolaire est très différent des pays du nord, pour éviter l'influence des facteurs culturels associés à ces différences de niveaux de développement. Nous n'avons pas pris en compte d'autres systèmes éducatifs, d'Asie

ou d'Océanie notamment, entre autres du fait de la nécessité de pouvoir donner sens à nos résultats en les rapportant à des contextes que nous connaissons et maîtrisons en partie.

Notre recherche porte donc sur un échantillon de 59 systèmes d'éducation. Elle intègre les systèmes éducatifs des pays membres de l'Union européenne (UE), de la Suisse, de l'Islande, de la Norvège non membres de l'UE, ainsi que ceux des provinces canadiennes. Puisque notre unité d'analyse est le système d'éducation, nous avons désagrégé les données produites par PISA au niveau pays, lorsque permis par la base de données. Les systèmes éducatifs de notre échantillon, les effectifs d'élèves sont présentés en annexe pour ne pas alourdir cette sous-section (ANNEXE 2).

Nous avons fait le choix de distinguer les régions espagnoles du fait de la forte autonomie dont ces dernières bénéficient dans la gestion de l'éducation. Seulement six des juridictions allemandes (länder) sont incluses à notre échantillon car dans certaines, le nombre d'élèves s'est révélé insuffisant³². Nous n'avons pas pu travailler à partir des données désagrégées pour la Suisse. Les données disponibles pour les cantons suisses faisant partie d'un projet spécial mené par l'OCDE, la comparabilité des résultats des élèves avec le reste de notre échantillon s'est avérée problématique. Finalement, nous aurions souhaité intégrer les systèmes éducatifs états-uniens à notre analyse, ces derniers offrant des repères précieux pour situer les SABP et leurs instruments. Néanmoins, les données désagrégées au niveau du système d'éducation ne sont pas accessibles pour les États-Unis pour le PISA 2012.

3.4 Démarche méthodologique et méthodes d'analyse des données : les deux volets empiriques de la recherche

Le cadre méthodologique général de notre recherche étant posé, nous présentons dans les sous-sections suivantes la démarche méthodologique, l'opérationnalisation des concepts clefs de notre étude, les variables mobilisées et les méthodes d'analyse des données.

La présentation est structurée autour des deux volets empiriques de notre recherche. Ils sont déployés pour répondre à la question centrale des effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation. Ces volets de recherche, bien que complémentaires,

³² Malgré nos efforts pour obtenir les données désagrégées pour les länder allemands auprès de l'institut allemand en charge de ces dernières — Institut zur Qualitätsentwicklung im Bildungswesen — nous n'avons pu les obtenir puisqu'elles n'étaient pas disponibles jusqu'il y a peu pour l'édition du PISA 2012.

poursuivent des objectifs et s'appuient sur des méthodes d'analyse des données distinctes. Le tableau synoptique ci-dessous (Tableau 3) en propose une synthèse.

Tableau 3 : Tableau synoptique du design de recherche

	Volet empirique 1	Volet empirique 2
Question centrale	Quels sont les effets des SABP sur le degré d'efficacité et d'équité scolaires des systèmes d'éducation ?	
Sous-questions de recherche	Quelles formes revêtent les SABP dans les systèmes d'éducation d'Europe et du Canada ?	Les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ?
Objectifs	<u>Descriptif et analytique</u> : 1) caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada ; 2) classifier les systèmes d'éducation de l'échantillon sur cette base <u>Opérationnel</u> : 3) construction de variables permettant d'analyser les effets des SABP sur l'efficacité et l'équité scolaires	<u>Analytique</u> : 1) analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires
Postulats/hypothèses	1) Les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes. 2) Les instruments permettent de différencier des types empiriques de SABP.	2) Effets différenciés des SABP (composantes, types empiriques) sur l'efficacité et l'équité scolaires
Concepts clefs	Modes de régulation institutionnelle, SABP	SABP, efficacité scolaire, équité scolaire (égalité de résultats, égalité des chances de résultats et égalité des acquis de base)
Méthodes d'analyse des données	Analyses multivariées : analyses en composantes principales, analyse de classification hiérarchique	Analyses bivariées (corrélations, régression simple, analyse de variance) Analyses multivariées (régressions multiples)

3.4.1 Premier volet empirique de recherche : démarche méthodologique et méthodes d'analyses des données

Les sous-sections suivantes reprennent les postulats et les objectifs au cœur de ce premier volet empirique de recherche. Les concepts clefs, leur opérationnalisation ainsi que les variables centrales y sont présentés. Les méthodes d'analyses mobilisées et leur pertinence y sont aussi discutées. À toutes fins utiles, le Tableau 4 reprend les éléments centraux de ce volet de recherche, déjà présentés dans le tableau précédent.

Tableau 4 : Tableau synoptique du volet empirique 1 de recherche

	Volet empirique 1
Questionnement central de la thèse	Quels sont les effets des SABP sur le degré d'efficacité et d'équité scolaires des systèmes d'éducation ?
Sous-question de recherche	Quelles formes revêtent les SABP dans les systèmes d'éducation d'Europe et du Canada ?
Objectifs	<u>Descriptif et analytique</u> : 1) caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada à partir de la gamme et de l'étendue des modes de régulation mis en œuvre au sein des établissements scolaires; 2) classier les systèmes d'éducation de l'échantillon sur cette base <u>Opérationnel</u> : 3) construction de variables permettant d'analyser les effets des SABP sur l'efficacité et l'équité scolaires.
Postulats	1) Les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes. 2) Les instruments permettent de différencier des types empiriques de SABP.
Concepts clefs	Modes de régulation institutionnelle, SABP
Méthodes d'analyse des données	Analyses multivariées : analyses en composantes principales, analyse de classification hiérarchique

3.4.1.1 Postulats et objectifs

Ce premier volet constitue la première étape empirique nécessaire à l'analyse des effets des SABP sur les degrés d'efficacité et d'équité scolaires. Il se focalise plus particulièrement sur le concept de SABP et sur sa définition opératoire. Il trouve sa justification dans un manque dans la littérature en ce qui a trait à la typification empirique des SABP dans les systèmes éducatifs de notre échantillon. Les objectifs poursuivis sont donc multiples.

Il s'agit 1) d'analyser et de caractériser les SABP dans les systèmes d'éducation d'Europe et du Canada ; 2) de classier ces derniers sur cette base ; 3) de créer des variables permettant, dans un

second volet de recherche et d'analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires.

Plusieurs postulats en lien avec notre cadrage théorique, et en cohérence avec une entrée d'analyse par les modes de régulation institutionnelle, sont donc formulés : 1) les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes ; 2) ils permettent de différencier des types empiriques de SABP.

3.4.1.2 Opérationnalisation du concept de systèmes d'*accountability* basés sur la performance et variables mobilisées

Notre cadrage théorique nous a en effet permis de situer les SABP au travers des concepts de régulation et de modes de régulation institutionnelle. Si les modes de régulation institutionnelle réfèrent à la façon « dont les autorités publiques exercent la coordination, le contrôle et l'influence sur le secteur éducatif » (Barroso, 2004, p. 92 cité dans Maroy, 2006), nous avons soutenu que les SABP peuvent être analysés comme une expression particulière de la « nouvelle régulation » des systèmes d'éducation et qu'ils reposent sur un ensemble d'instruments qui visent à orienter et à contrôler les comportements des acteurs du champ éducatif.

Ce volet de recherche s'inscrit ainsi dans une conception large de l'*accountability* et vise à caractériser les SABP à partir de la gamme d'instruments mis en œuvre au sein des établissements scolaires. Il ne s'agit donc pas de tester l'efficacité d'une théorie de l'action³³ et de cibler les instruments qui s'y rapportent tel que l'ont fait les études quantitatives internationales jusqu'à lors (voir en particulier Wößmann, 2007 ; Schütz, West & Wößmann, 2007). Il s'agit de prendre en compte les instruments effectivement prégnants au sein des systèmes d'éducatifs, dont la gamme et l'étendue peuvent être sensiblement diversifiés en fonction des contextes, qui constituent le socle des SABP. Aussi des méthodes d'analyse permettant de faire émerger de nos données d'une part des logiques de régulation distinctes, d'autre part des types empiriques de SABP ont été privilégiées (nous y reviendrons).

³³ Tel que nous l'avons mis en exergue dans notre revue de littérature, Wößmann (2007) teste finalement l'efficacité de l'*accountability* de marché qui repose sur trois éléments centraux : la compétition et la concurrence entre établissements scolaires, l'autonomie scolaire et l'*accountability*. Cette dernière fait référence d'une part à la présence d'examens standardisés permettant d'évaluer la capacité des établissements scolaires à favoriser les apprentissages effectifs des élèves, d'autre part, aux mécanismes de reddition de comptes au public (publication des résultats des élèves) permettant en outre de stimuler la compétition entre établissements.

Dès lors, nous posons qu'une entrée d'analyse par l'instrumentation, qui sous-tend et met en acte les SABP, s'avère pertinente pour pouvoir caractériser ces derniers et en analyser les effets sur l'efficacité et l'équité scolaires. Nous nous proposons donc d'opérationnaliser les concepts de SABP en fonction d'une série de variables de la base de données du PISA 2012 se rapportant à cette instrumentation.

Nous avons donc sélectionné 17 variables dans la base de données « écoles » du PISA 2012. Elles nous renseignent sur les instruments mobilisés au sein des systèmes éducatifs, dans une perspective de régulation par les résultats.

Elles sont regroupées dans la base de données du PISA selon quatre séries ayant trait 1) aux dispositifs ou aux instruments mobilisés par les établissements aux fins d'assurance qualité et d'amélioration des écoles 2) au suivi des résultats des écoles (suivi par une autorité administrative et publication des résultats) 3) à l'usage des résultats des élèves par les écoles pour différents propos 4) aux incitatifs ou aux conséquences articulés aux évaluations des enseignants.

Ces variables sont présentées dans le Tableau 5. Les questions adressées aux chefs d'établissements sont reprises et elles sont laissées en anglais. Nous assumons par contre la traduction des intitulés de réponses sélectionnées pour les besoins de notre recherche. Néanmoins, le lecteur peut se référer aux questionnaires placés en annexe (ANNEXE 4).

Tableau 5 : Opérationnalisation du concept de système d'*accountability* basés sur la performance, variables mobilisées (PISA 2012)

VARIABLES DU PISA 2012 —INSTRUMENTS DES SABP ³⁴			
Série 1	Série 2	Série 3	Série 4
Dispositifs ou instruments mobilisés dans les établissements aux fins d'assurance qualité et d'amélioration des écoles	Suivi des performances des écoles et publication des résultats des élèves	Usage des résultats des élèves au niveau des établissements scolaires	Suites, conséquences (incitatifs) des évaluations des enseignants
<i>Which of the following measures aimed at quality assurance and improvement do you have in your school?</i> (SC 39)	<i>In your school, are achievement data used in any of the following <accountability procedures</i> (SC 19)	<i>In your school, are assessments of students in <national modal grade for 15-year-olds> used for any of the following purposes?</i> (SC 18)	<i>To what extent have appraisals of and/or feedback to teachers directly led to the following?</i> (SC 31)
Consultation d'experts dans les six derniers mois (Q09)	Suivi par une autorité administrative (Q02)	Comparaison des performances de l'école à celles des autres écoles (Q08)	Attribution de responsabilités professionnelles (Q06)
Évaluation externe (Q06)	Publication des résultats (Q01)	Évaluation de la performance des enseignants (Q06)	Reconnaissance publique de la part du chef d'établissement (Q05)
Évaluation interne (Q05)		<i>Monitoring</i> des performances des écoles d'une année sur l'autre (Q05)	Implications en termes d'avancement de carrière (Q04)
Enregistrement systématique des données incluant les taux de présence et de graduation des enseignants et des élèves, résultats aux tests, participation des enseignants à des activités de développement professionnel (Q03)		Comparaison des performances de l'école aux performances régionales ou nationales (Q04)	Activités de développement professionnel (Q03)
Spécification des standards de performance à atteindre (Q02)			Bonus financier (Q02)
			Changement de salaire (Q01)

³⁴ Les références SC39, SC19, SC18, SC31 identifient les questions du questionnaire école du PISA 2012, les références Q0 indique l'identification des variables de la série telle que reporté dans le « *code book* » PISA).

La première série (1) de variables renseigne sur les dispositifs mobilisés par les écoles aux fins d'amélioration et d'assurance qualité. Nous considérons qu'elles permettent de saisir les supports, les mécanismes et les dispositifs mis en œuvre au sein des établissements scolaires qui devraient théoriquement permettre d'améliorer les performances des élèves et des écoles.

La seconde série (2) renvoie, comme le PISA l'indique, aux mécanismes de reddition de comptes. Ils ont trait d'une part au suivi des performances des élèves par une autorité administrative et d'autre part, à la publication des résultats des élèves.

La troisième série (3) nous permet de saisir les usages des résultats des élèves par les écoles, à la fois dans une perspective de micro-pilotage et dans une dynamique de compétition entre écoles, pour évaluer la performance des enseignants ou pour comparer les performances des écoles aux performances régionales et/ou nationales.

Enfin, la dernière série (4) permet de saisir les conséquences associées aux évaluations des enseignants, qui prennent diverses formes. Autrement dit, selon les contextes, quels sont les suites ou les enjeux des évaluations des enseignants³⁵ ?

Ces variables sont agrégées au niveau du système éducatif en respectant les procédures recommandées par l'OCDE³⁶ (OCDE, 2009). Les variables construites rendent compte de la proportion d'élèves scolarisés dans les écoles où le chef d'établissement a déclaré ces instruments en usage (modalité de réponses oui, pour les questions SC39, SC19, SC18, SC31). Elles prennent la forme de variables continues³⁷.

³⁵ À noter ici que les réponses à cette question ne permettent pas d'apprécier le caractère positif ou négatif des incitatifs associés. En d'autres termes, s'agissant par exemple du changement de salaire, il ne nous est pas possible de savoir si le changement de salaire associé aux évaluations des enseignants prend la forme d'une augmentation ou d'une diminution. Aussi, la base de données ne nous renseigne pas sur les « sanctions négatives » qui pourraient être associées aux évaluations des enseignants telles que le licenciement ou le changement d'établissement, prévus par exemple dans certains contextes.

³⁶ L'agrégation des informations au niveau du système d'éducation est réalisée sur base des recommandations de l'OCDE en termes de pondérations du poids des effectifs dans chaque système d'éducation (OCDE, 2009). Ces procédures d'agrégation sont réalisées à partir d'un module SPSS (SPSS replicates) qui nous a été fourni par FORS Suisse — *Suisse Center for Expertise in Social Sciences* — et crée par l'*Australian Council for Education Research* (ACER). Le choix d'agréger les proportions d'élèves scolarisés dans des établissements où certains instruments sont déclarés en usage (plutôt que les proportions d'établissements) repose sur l'observation des procédures mobilisées par l'OCDE dans ces rapports.

³⁷ Dans la base de données originale, les trois premières séries de variables (SC39, SC19, SC18) se présentent sous une forme dichotomique (oui/non). La dernière série de variables (SC31) propose une gradation concernant les effets des évaluations des enseignants sur différentes dimensions – attribution de bonus financier, changement de salaire, etc... Par l'échelle proposée, le PISA tente donc de saisir l'ampleur des changements qui font suite, ou non, à l'évaluation des enseignants. Nous recodons donc cette variable qui propose à l'origine plusieurs modalités de

Les statistiques descriptives réalisées à partir de ces variables sont placées en annexe (ANNEXE 5). Elles permettent de mettre en exergue l'importante variation entre les systèmes d'éducation en ce qui a trait à l'usage de ces instruments. Ces dernières nous renseignent donc sur un ensemble de modes de régulation que nous considérons au cœur des SABP. Ces variables sont au centre des analyses que nous présentons dans la sous-section suivante.

Tel que nous le préciserons dans les limites de notre travail, certaines variables pertinentes ne sont pas disponibles dans la base de données du PISA 2012. En premier chef, la fréquence des évaluations standardisées des élèves pourtant fournie dans l'édition 2009. Même si nous avons pu constater que ce type d'évaluation était en œuvre dans la majorité des systèmes éducatifs de notre échantillon, des informations complémentaires auraient été nécessaires du fait de l'importance jouée par ce dispositif dans les SABP. Par ailleurs, la littérature a montré que les évaluations externes standardisées s'avèrent clefs puisque potentiellement en lien avec les degrés d'efficacité et d'équité scolaires (voir entre autres à ce sujet Mons, 2009). L'usage des taux de performance pour évaluer les performances des chefs d'établissement, aussi fournie par le PISA 2009, aurait aussi été fort pertinent pour nous renseigner sur les enjeux attachés aux résultats des élèves pour d'autres catégories d'acteurs.

3.4.1.3 Méthodes d'analyse des données

Sur base de ces données agrégées, nous réalisons deux séries d'analyses : des analyses en composantes principales (ACP) et des analyses de classification hiérarchique.

Ces analyses présentent plusieurs intérêts : d'abord descriptif et analytique, puisqu'elles permettent, à partir des instruments de régulation mis en œuvre au sein des établissements scolaires, de mettre en exergue, pour le cas des ACP, les configurations instrumentales privilégiées et les logiques de régulation qui les sous-tendent, à partir des associations statistiques entre nos variables d'intérêt. D'autre part, les analyses de classification hiérarchique permettent d'élaborer une typologie empirique des SABP, en d'autres termes, de classer les

réponse – 1) aucun changement (*no change*), 2) petit changement (*small change*), 3) changement modéré (*moderate change*), 4) changement important (*large change*). Les modalités 2 à 4 sont regroupées. La variable recodée prend la forme d'une variable dichotomique oui/non à partir de laquelle les proportions d'élèves se rapportant à la modalité de réponse « oui » sont agrégées au niveau du système d'éducation. Ce qui nous intéresse en effet est d'observer si oui ou non, les évaluations des enseignants engagent des conséquences sur différentes dimensions renseignées par le PISA. Nous considérons donc que la perte d'information due au recodage n'est pas problématique pour notre propos.

systèmes d'éducation au sein de types de SABP, en fonction des configurations instrumentales qui y sont privilégiées. Finalement, ces analyses aboutissent à la création de variables mobilisées dans le second volet empirique de notre recherche pour analyser les effets des SABP sur l'efficacité et l'équité scolaires.

La pertinence de ces méthodes d'analyse étant posée, nous les décrivons en termes d'avantages techniques dans les sous-sections suivantes. Nous montrons par ailleurs les possibilités interprétatives offertes par ces méthodes d'analyses : mise en exergue des logiques instrumentales des SABP dans le cadre des ACP, création de types empiriques de SABP pour les ACH.

3.4.1.3.1 Analyse en composantes principales

Les analyses en composantes principales (ACP) permettent d'une part, nous l'avons déjà dit, de mettre en exergue les associations statistiques entre les variables mobilisées, d'autre part d'opérer une réduction dans la masse des données en vue des analyses de classification hiérarchique. Elles permettent finalement d'aboutir à la construction d'une première série de variables mobilisées dans le deuxième volet empirique de notre recherche.

En ce qui a trait à la mise en exergue des associations statistiques entre nos variables d'intérêt, rappelons que l'ACP consiste à transformer des variables corrélées entre elles³⁸ en composantes indépendantes les unes des autres, en d'autres termes, en axes indépendants expliquant au mieux la variabilité des données (Bouroche & Saporta, 1980). On souhaite ainsi résumer l'information contenue dans l'ensemble des variables (la variance totale) sur un nombre limité de dimensions (les composantes principales) qui réalisent la synthèse des caractères initiaux au moyen de combinaison linéaire (Bourque, Poulin, & Cleaver, 2006).

L'ACP vise ainsi « à réduire le nombre de variables requises pour calculer un score total pour le trait mesuré, tout en conservant le plus d'informations possible de l'ensemble original » (Bryant & Yarnold, 1995; Fabrigar *et al.*, 1999 cités dans Bourque, Poulin, Cleaver, 2006, p. 326) :

“The goal of PCA (e.g Principal Component Analysis) is simply to reduce the number of variable by creating linear combinations that retain as much of the original measures' variance as possible” (Conway & Huffcutt, 2003, p. 150).

³⁸ Le coefficient de corrélation linéaire est à la base de la mesure d'intensité de la relation entre deux variables qui dans ce cadre doivent être quantitatives.

Il faut ajouter au propos de Conway et Huffcutt (2003) que les ACP ont un autre intérêt que la simple réduction des variables. En effet, pour reprendre les propos de Bouroche et Sapporta « elles dévoilent des relations entre caractères » (Bouroche & Sapporta, 1980, p. 36). Ainsi, les composantes extraites représentent un trait, une dimension qu'il s'agit d'interpréter théoriquement. Il s'agit de donner un sens aux composantes extraites, qui constituent donc des « caractères principaux (ou composantes principales) réalisant la synthèse de plusieurs caractères initiaux au moyen d'une combinaison linéaire de ces caractères initiaux » (Bouroche & Sapporta, 1980, p. 30).

Précisons d'entrée de jeu que si, à la lumière de certains écrits (Everitt, 1975; Gorsuch, 1983), notre ratio cas/variables de 3,47 (59/17) peut paraître faible pour réaliser des ACP, il est toutefois considéré comme acceptable si nous nous basons sur les travaux de Cattell (1978). Ce dernier préconise un ratio de 3 à 5 variables par cas. En outre, pour Mac Callum, Widaman, Zang et Hong (1999), la taille de l'échantillon et le ratio cas/variables sont des critères insuffisants pour juger de la qualité de la solution factorielle. Pour ces auteurs, ce sont d'autres paramètres centraux que les chercheurs doivent analyser : le degré de saturation des variables sur les composantes (autrement appelé *communalités*) qui doit être élevé (valeurs supérieures à 0,6), le niveau de surdétermination des facteurs, en d'autres termes, le nombre de variables construisant les composantes (environ six indicateurs par facteur et un nombre restreint de facteurs émergeant de l'analyse) conditionnant la capacité du chercheur à interpréter théoriquement les facteurs. Ces critères étant respectés, les résultats de nos analyses en composantes principales pourront donc être robustes.

Nous réalisons une ACP avec rotation orthogonale (varimax) qui permet d'extraire des composantes indépendantes de la masse de nos données, c'est-à-dire non corrélées les unes aux autres. Outre le fait que cette rotation simplifie l'interprétation des composantes et des variables associées à ces dernières (Tabachnick & Fidell, 2013), le choix d'une rotation orthogonale est justifié d'une part, par le fait que les composantes créées via les ACP (et donc les variables construites – voir infra) seront intégrées à des analyses de régressions multiples. L'absence de corrélation entre les composantes permet donc de diminuer les problèmes de colinéarité. D'autre part, si, empiriquement, différentes dimensions au cœur des SABP sont inter-reliées, nous souhaitons, dans notre second volet de recherche, analyser l'impact potentiellement différencié de ces composantes sur l'efficacité et l'équité scolaires. Une rotation orthogonale qui permet

d'extraire l'information unique et non partagée par les autres composantes apparaît donc plus pertinente pour notre propos.

Les composantes à retenir et la qualité de la solution factorielle sont analysées sur base des critères établis dans la littérature (Baillargeon, 2003; Hair, Black, Babin, & Anderson, 2010; Ho; Tabachnick & Fidell, 2013). Le nombre de composantes à retenir est obtenu sur la base du diagnostic des valeurs propres (*eigenvalue* supérieure à 1). Un ensemble de paramètres permettant de juger de la qualité de la solution factorielle est aussi observé : inertie totale et part de variance expliquée par les composantes, indice de Kaiser-Mayer-Olkin, significativité du test de sphéricité de Bartlett.

Nous réalisons une série de tests pour mesurer la cohérence interne des composantes. Il s'agit donc de vérifier statistiquement si les variables associées aux composantes mesurent de façon constante le trait extrait (Ho, 2006). Ici, nous mobilisons l'alpha de Cronbach³⁹ pour tester la cohérence interne des composantes. Ho (2006) suggère une mesure de l'alpha de Cronbach supérieur à 0,80 pour attester de la solution retenue, mais une valeur supérieure à 0,70 est aussi considérée acceptable (Frisbie, 1988).

Finalement, l'interprétation des composantes repose sur l'observation des variables qui participent le plus fortement à la construction des composantes, ce qui se traduit, en termes statistiques, par les variables les plus fortement corrélées aux composantes extraites. Nous observons d'une part les variables qui construisent les composantes, ce à quoi McCallum, Widaman, Zang et Hong (1999) font référence sous le terme de détermination des facteurs et, d'autre part, les indices de saturation des variables sur ces dernières. Les ACP nous permettent ainsi d'observer comment différentes variables se regroupent et s'associent pour composer ces caractères.

En termes d'interprétation, ces variables éclairent donc les configurations d'instruments qui sous-tendent différentes composantes des SABP et qui sont effectivement mis en place par les établissements scolaires dans les différents systèmes éducatifs. Nous pourrions ainsi faire le lien avec notre cadrage théorique et les travaux de certains auteurs (Lascoumes & Le Galès, 2004 ;

³⁹ Certains auteurs ont toutefois pointé que d'autres tests peuvent être mobilisés comme alternative à l'Alpha de Cronbach ; cependant, cette mesure reste parmi les plus mobilisés dans des cas similaires au nôtre (Béland, Cousineau, & Loye, 2016).

Lascoumes & Simard, 2011 ; Maroy & Voisin, 2014) qui avancent que les instruments sont porteurs de logiques de régulation. Nous faisons en effet le postulat que des instruments différenciés sous-tendent finalement des logiques de régulation distinctes.

Finalement, pour chaque composante extraite, un score factoriel est attribué aux systèmes éducatifs de notre échantillon⁴⁰. La définition des scores factoriels est résumée de la sorte par Barroso Da Costa (2014) :

« Un score factoriel est le produit d'un modèle formatif composé de variables qui, bien qu'elles soient corrélées, sont indépendantes à l'origine (Bollen et Lennox, 1991). Le score factoriel (aussi nommé score composite) est donc le sommaire de la variation commune entre divers indicateurs. C'est l'estimation que chaque cas aurait reçue sur chacun des facteurs s'ils avaient été mesurés directement (Tabachnick et Fidell, 2007) » (Barroso Da Costa, 2014, p. 110).

Les scores factoriels nous permettent d'apprécier l'ampleur, en d'autres termes, de quantifier chaque composante extraite des systèmes éducatifs de notre échantillon, et de comparer ces derniers sur cette base. Ils constituent une série de variables (variable continue - échelle d'intervalle [(Bressoux, 2010)]) mobilisées dans le second volet empirique de notre recherche pour apprécier les effets des composantes instrumentales des SABP sur l'efficacité et l'équité scolaires des systèmes éducatifs de notre échantillon. Ces scores factoriels constituent la base des analyses de classification hiérarchique que nous présentons dans la sous-section suivante.

3.4.1.3.2 Analyse typologique : analyse de classification hiérarchique

L'analyse de classification hiérarchique, autrement nommée analyse de *cluster*, est la deuxième série d'analyses mobilisée dans ce premier volet de recherche. Cette méthode d'analyse est utilisée comme support à l'élaboration d'une typologie empirique des SABP. Cette méthode essentiellement descriptive (Field, 2009) nous permet d'observer la construction de types empiriques de SABP où différentes logiques instrumentales de régulation (extraites des ACP) sous-tendues par des instruments distincts, se combinent et s'articulent. Les types de SABP dégagés permettent ainsi de voir les similitudes entre systèmes éducatifs regroupés sous un même type en ce qui a trait aux configurations instrumentales privilégiées, mais aussi les différences entre systèmes d'éducation appartenant à des types de SABP différents.

⁴⁰ La méthode « régression » est utilisée pour la création des scores factoriels.

L'analyse typologique de type *cluster* revêt en outre un caractère exploratoire et descriptif. Elle permet de dresser une première classification empirique des SABP dans les systèmes d'éducation d'Europe et du Canada.

Cette démarche d'analyse se rapporte donc au deuxième postulat au cœur de ce premier volet de recherche : l'articulation des composantes et donc, différentes configurations d'instruments, permettent de distinguer différents types empiriques de SABP privilégiés au sein des systèmes éducatifs de notre échantillon.

Selon Falissard (1996) l'analyse typologique vise à « découper un jeu de données en sous-groupes homogènes (...) ce découpage est effectué *de novo* : les *clusters* ne sont pas définis à partir d'une variable externe (comme dans le cas de l'analyse discriminante), mais à partir de la structure même des données » (Falissard, 1996, p. 185). Les potentialités de l'analyse et ses enjeux sont illustrés par les propos de Field (2009) :

“Cluster analysis is an approach that finds structure in data by identifying natural grouping (clusters) in the data. Unfortunately ‘natural groupings’ is not as well defined as we might hope. Indeed, it is usual to have more than one natural grouping for any collection of data. As we will see there is no definitive clusters analysis technique, instead the terms relate to a rather loose collection of algorithms that groups similar objects into categories (clusters). Although some clustering algorithms have been present in ‘standard’ statistical software for many years, they are rarely used for formal significance testing. Instead they should be viewed as EDA (eg exploratory data analyses) tools because they are generally used to generate, rather than test, hypotheses about data structure” (Field, 2009, p. 47)

La citation empruntée à Field (2009) permet par ailleurs d'introduire le fait que l'analyse typologique repose sur une série de choix techniques laissés à la discrétion du chercheur. Les choix que nous avons effectués se basent principalement sur l'analyse des travaux sur la question, dont ceux de Rapkin et Luke (1993) et de Everitt, Landau, Leese et Stahl (2011).

Nous avons choisi d'employer une méthode d'analyse de classification hiérarchique de type agglomératif. Le choix d'utiliser cette méthode plutôt qu'une analyse de type nuée dynamique ou hiérarchique divisif a trait essentiellement au nombre de cas relativement restreint dans notre étude. Dans ce cas de figure, cette méthode de classification est considérée comme la plus appropriée (Rapkin & Luke, 1993). Nous avons utilisé comme mesure de distance la « distance euclidienne au carré », l'algorithme de regroupement Ward (Everitt, Landau, Leese & Stahl, 2011). Cette mesure de distance et l'algorithme choisis permettent de minimiser la variation

intra-classe et de maximiser la distance inter-*cluster*. L'algorithme de regroupement de Ward, dit compact, regroupe les cas (ici, les systèmes d'éducation) de sorte que chaque cas est plus similaire à tous les membres de sa classe qu'aux membres des autres classes (Blashfield & Aldenderfer, 1988; Rapkin & Luke, 1993). De fait, ils visent à mettre en évidence les différences inter-classes (Saint-Arnaud, 2001). L'association entre la distance euclidienne au carré comme mesure de distance et l'algorithme Ward est, en outre, recommandé dans la littérature (Everitt, Landau, Leese & Stahl, 2011).

L'analyse de classification hiérarchique aboutit à la définition d'un ensemble de classes illustrées dans un dendrogramme. Ici, le nombre de classes qu'il est pertinent de choisir est aussi laissé à la discrétion du chercheur. Ce choix repose d'une part sur l'examen du graphique, mais aussi sur la connaissance de son objet d'étude que ce dernier a pu développer. Par ailleurs, lorsque ces catégories ont vocation à être mobilisées dans des analyses subséquentes, la littérature suggère une répartition relativement équilibrée des cas au sein des types construits.

Aux fins de contrôle, nous avons comparé les résultats d'analyses de classification hiérarchique mobilisant différentes mesures de distances ainsi que plusieurs algorithmes de regroupement. Puisqu'il n'existe pas de solutions idéales et que la pertinence de la solution reste soumise au jugement du chercheur, cette démarche nous a permis d'observer dans quelle mesure certains types empiriques (ou de classes, deux termes que nous utiliserons dans notre travail) se dessinent et le fait que la solution reste stable malgré une variation des paramètres clefs de l'analyse de classification. Au final, l'ensemble des procédures d'analyse aboutit à des solutions relativement similaires, aussi bien au niveau du nombre de classes qu'il est pertinent de choisir, que des regroupements de systèmes éducatifs qui sont effectués.

Il est par ailleurs important de rappeler un point d'ordre méthodologique ayant des implications sur les résultats de l'analyse de classification et donc, sur la « réalité » empirique organisée. Comme nous l'avons observé dans la partie précédente, la contribution des variables initiales à la formation des composantes extraites de l'ACP n'est pas égale d'une variable à l'autre, ce qui implique que certaines variables ont un poids plus important que d'autres dans la construction des composantes. La même logique de contribution différenciée des composantes s'applique à l'ACH. De fait, certaines composantes ont un poids plus important que d'autres dans la construction des types. Pour mettre en exergue la contribution différenciée des composantes dans

la construction des types et faciliter l'interprétation des types empiriques, nous réalisons une analyse de variance sur les composantes principales.

Au total, l'analyse de classification hiérarchique permet la construction d'une typologie empirique des SABP où sont regroupés les systèmes éducatifs de notre échantillon. Cette typologie donne lieu à la construction d'une variable nominale (échelle nominale [Bressoux, 2010] à plusieurs modalités) qui sera mobilisée dans le deuxième volet de recherche empirique pour analyser les effets potentiellement différenciés de ces types sur l'efficacité et l'équité scolaires.

En résumé, nous mobilisons dans ce volet de recherche deux types d'analyse. Au-delà de leur intérêt descriptif, ces analyses nous permettent de construire deux séries de variables permettant de caractériser les SABP dans les systèmes éducatifs de notre échantillon et de classer ces derniers sur cette base. Les variables mobilisées et les variables créées dans ce volet de recherche sont synthétisées dans le Tableau 6.

Tableau 6 : Opérationnalisation du concept de système d'*accountability* basé sur la performance, méthodes d'analyses des données et variables mobilisées

Méthodes d'analyses des données	Analyses en composantes principales [ACP]	Analyse de classification hiérarchique [<i>Cluster</i>]
Variables mobilisées [PISA 2012]	1) Dispositifs et instruments aux fins d'assurance qualité et d'amélioration des écoles (SC39) 2) Mécanismes de reddition de comptes (SC19) 3) Usage des résultats des élèves par les établissements scolaires (SC18) 4) Suites, conséquences des évaluations des enseignants (SC31)	Composantes instrumentales (scores factoriels attribués aux systèmes d'éducation de l'échantillon) créées via les analyses en composantes principales réalisées sur les variables présentées ci-contre
Variables créées	Composantes instrumentales des SABP	Types d'instrumentation des SABP selon les systèmes éducatifs
Types de variable créée	Variable continue — échelle d'intervalle (scores factoriels)	Variable nominale — échelle nominale à quatre modalités

Ces méthodes d'analyse liées à notre premier volet de recherche empirique étant précisées, nous présentons, dans les sous-sections suivantes, la démarche méthodologique qui entoure le second volet empirique de notre recherche. Nous précisons comment les variables créées dans ce premier volet empirique seront mobilisées dans des analyses subséquentes.

3.4.2 Second volet empirique de recherche : démarche méthodologique et méthodes d'analyses des données

Ce second volet de recherche est fortement articulé au premier volet de recherche qui aboutit à la construction de variables permettant de caractériser différentes composantes et types empiriques de SABP. Il vise finalement à analyser et à discuter les effets des SABP sur les degrés d'efficacité et d'équité des systèmes éducatifs de notre échantillon. Plusieurs types d'analyses seront donc mobilisés et permettront de répondre à la question centrale de notre recherche doctorale.

Les sous-sections suivantes reprennent les hypothèses au cœur de ce second volet empirique de recherche. Les concepts clefs et leur opérationnalisation y sont présentés. Les méthodes d'analyses mobilisées et leur pertinence y sont aussi discutées.

À toutes fins utiles, le tableau ci-dessous reprend certains des éléments clefs de ce volet de recherche.

Tableau 7 : Tableau synoptique du volet empirique 2 de recherche

	Volet empirique 2
Questionnement central de la thèse	Quels sont les effets des SABP sur le degré d'efficacité et d'équité scolaires des systèmes d'éducation ?
Sous-question de recherche	Les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ?
Objectifs	<u>Analytique</u> : 1) analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires
Hypothèse	2) Effets différenciés des SABP (composantes, types empiriques) sur l'efficacité et l'équité scolaires (égalité de résultats, égalité des chances de résultats et égalité des acquis de base)
Concepts clefs	SABP, efficacité scolaire, équité scolaire
Principales méthodes d'analyse des données	Analyses bivariées (corrélations, régression simple, analyse de variance) Analyses multivariées (régressions multiples)

3.4.2.1 Hypothèses et objectifs

Ce second volet de recherche vise à produire des évidences empiriques permettant de répondre à la question centrale des effets des SABP sur l'efficacité et l'équité scolaires. L'hypothèse est que les SABP influencent de manière différenciée les degrés d'efficacité et d'équité scolaires en fonction des composantes instrumentales privilégiées et des types empiriques qui sont construits sur cette base.

Rappelons ici que le premier volet de recherche a permis d'aboutir à la création de deux types de variables permettant de caractériser les SABP au sein de notre échantillon. La première, les scores factoriels attribués aux systèmes d'éducation, prend la forme d'une variable continue (échelle d'intervalle) permettant de mesurer l'ampleur de chaque composante instrumentale des SABP (extraite des ACP) au sein des systèmes éducatifs de notre échantillon. La seconde prend la forme d'une variable nominale (échelle nominale) à plusieurs modalités où sont regroupés les systèmes éducatifs de notre échantillon. Elle permet de caractériser des types d'instrumentation des SABP selon les systèmes éducatifs. Ces variables seront donc mises en lien dans nos analyses (variables dépendantes [X]) avec les indicateurs d'efficacité et d'équité scolaires (variables dépendantes [Y]) que nous présentons dans les sous-sections suivantes.

3.4.2.2 Opérationnalisation des concepts d'efficacité et d'équité scolaires : indicateurs et variables mobilisées

Pour rappel, comme discuté en amont, l'un des enjeux phares au cœur du développement des SABP dans les systèmes d'éducation est l'amélioration de la qualité de l'éducation appréhendée au travers des notions clefs d'efficacité et d'équité scolaires. Si le débat qui entoure la préoccupation grandissante d'efficacité des systèmes éducatifs tend à se centrer sur la question de l'augmentation des performances des élèves, les questions d'équité scolaire renvoient quant à elles au jugement normatif qui est posé quant au statut des inégalités de performance et d'acquisition scolaires au sein des systèmes d'éducation. Puisque nous souhaitons analyser les effets des SABP sur ces deux dimensions, nous opérationnalisons ces concepts clefs en une série d'indicateurs que nous présentons dans les sous-sections suivantes.

3.4.2.2.1 Degré d'efficacité scolaire

Le concept d'efficacité scolaire réfère à la capacité des systèmes d'éducation à atteindre les objectifs visés. Il sera opérationnalisé en référence aux résultats des élèves aux épreuves du PISA 2012. Précisons à nouveau que les tests du PISA ne sont pas alignés sur les standards académiques et de performance des systèmes éducatifs (et vice versa). Néanmoins, ces épreuves permettent de définir un ensemble de connaissances et de compétences définies comme clefs par l'OCDE. Puisque le nombre de pays participants ne cesse d'augmenter, on peut considérer que cette participation a valeur de consensus au sein des économies participantes, au moins en théorie, en ce qui a trait à l'importance accordée aux acquisitions évaluées par le PISA.

En cohérence, l'indicateur choisi pour rendre compte du degré d'efficacité scolaire réfère aux performances moyennes des élèves en mathématiques⁴¹ dans les systèmes d'éducation de notre échantillon (moyennes des élèves agrégées au niveau système éducatif). Il est en outre mobilisé dans nombre d'études s'intéressant à l'efficacité des systèmes d'éducation (Mons, 2004; Le donnée, 2017).

Les variables de la base de données du PISA 2012 qui indiquent les résultats des élèves au test en mathématiques sont mobilisées. Précisons que la base de données du PISA « élèves » inclut cinq mesures de performance des élèves pour chaque discipline testée. La compétence des élèves est donc évaluée en fonction de cinq « valeurs plausibles⁴² » (PV 1 à 5) et non par un score unique :

“Student proficiency estimates are returned through plausible values. ‘The simplest way to describe plausible values is to say that plausible values are a representation of the range of abilities that a student might reasonably have. (...). Instead of directly estimating a student’s ability, a probability distribution for a student’s, is estimated. That is, instead of obtaining a point estimate for, (like a WLE4), a range of possible values for a student’s,

⁴¹ L'étude porte uniquement sur les mathématiques. Nous avons dans un premier temps souhaité mobiliser les trois matières, toutefois, des problèmes dans la distribution des données pour les sciences (indicateur d'efficacité scolaire et d'inégalité de résultats) et la lecture (indicateur d'inégalités de résultats). De fait, les résultats présentés portent uniquement sur les mathématiques.

⁴² Vriгдаud (2006) explicite la logique qui sous-tend les valeurs plausibles : « le score de compétence du sujet est conceptuellement une valeur non observée et que son estimation renvoie non pas à un seul paramètre, mais à une distribution. Conditionnellement aux réponses et aux caractéristiques de ce sujet, on infère, avec une plus ou moins bonne garantie, la distribution du paramètre de compétence d'un sujet ayant ces caractéristiques et ce patron de réponses aux items. On ne connaît pas la valeur vraie du paramètre de compétence, mais sa distribution. Pour renforcer la robustesse de cette estimation, on va procéder à plusieurs tirages dans cette distribution de valeurs dites plausibles dont la moyenne sera une meilleure estimation de la compétence de ce sujet ». Pour plus d'information nous renvoyons par ailleurs aux travaux de l'OCDE ayant visé à accompagner l'usage des données des enquêtes PISA (OCDE, 2009)

with an associated probability for each of these values is estimated. Plausible values are random draws from this (estimated) distribution for a student's”(Wu & Adams, 2002) » (OCDE, 2009, p. 43)

Si nous ne revenons pas sur les détails de cette approche (développée en psychométrie et d'ailleurs en débat dans ce champ) et les procédures sur lesquelles elle repose, précisons que les valeurs plausibles devraient théoriquement permettre un meilleur calcul des moyennes de performances des pays (Vrignaud, 2006).

Ce qui nous importe de savoir est que l'emploi des valeurs plausibles doit répondre à certaines procédures (explicitées par l'OCDE) pour les calculs de différents paramètres de distribution (moyenne, écarts-types, etc.). Nous les avons donc prises en compte pour l'ensemble des analyses mobilisant les scores des élèves⁴³. Pour l'indicateur d'efficacité scolaire, le calcul de la moyenne des scores est effectué dans chaque système éducatif.

3.4.2.2 Degré d'équité scolaire

Comme discuté dans notre cadrage théorique, s'intéresser à l'équité des systèmes éducatifs nécessite de définir différentes figures de l'égalité au prisme desquelles l'équité scolaire sera donc évaluée. Dans notre recherche doctorale, le degré d'équité scolaire est appréhendé au travers de plusieurs indicateurs permettant la mesure de trois types d'inégalités au sein des systèmes scolaires. Ils sont, comme l'indicateur d'efficacité scolaire, déclinés pour les mathématiques. Les variables du PISA 2012 permettant de rendre compte des résultats des élèves aux tests dans cette discipline sont ici mobilisées.

⁴³ Les analyses mobilisant les valeurs plausibles pour l'estimation des paramètres de distribution (définitions des moyennes, écarts-types, dispersion des scores, modèles de régression simple) et la constitution de variables agrégées au niveau du système éducatif ont été réalisées à partir du module SPSS fourni par FORS et développé par ACER qui permet de prendre en compte les procédures de réplification ainsi que les critères de pondération fortement suggérés par l'OCDE (OCDE, 2009). Les scores des élèves en mathématiques à partir desquels les indicateurs d'efficacité et d'équité scolaires sont donc calculés à partir des 5 valeurs plausibles. Pour plus de détail sur l'usage des valeurs plausibles et l'utilisation des procédures de réplification et d'agrégation des données le lecteur peut se référer au manuel d'utilisation des données du PISA (OCDE, 2009). Nous ne reviendrons donc pas sur cette précision dans la suite de notre travail.

Inégalité de résultats

L'égalité de résultats, et son corollaire l'inégalité de résultats, renvoie aux différences d'acquisitions des élèves au sein des systèmes d'éducation. L'indicateur d'inégalité de résultats que nous construisons vise à mesurer l'ampleur des différences de résultats entre les élèves les plus performants et les élèves les plus faibles au sein des systèmes d'éducation. Il est construit sur base des travaux de Dupriez et Dumay (2005), et renvoie au rapport entre le dernier (D9) et le premier décile (D1) de la distribution des scores (D9/D1) au sein de chaque système éducatif.

Inégalité des chances de résultats

La notion d'inégalité des chances de résultats réfère à la détermination socioculturelle des résultats scolaires, en d'autres termes, à l'influence du milieu socioéconomique sur les résultats scolaires. En termes statistiques, ce concept peut donc être opérationnalisé par la proportion de la variabilité des scores des élèves expliquée par leur milieu socioéconomique.

L'indicateur d'inégalité des chances de résultats réfère donc à la valeur du coefficient de détermination d'une régression linéaire où la variable dépendante (Y) est le score des élèves dans chaque matière évaluée par le PISA, la variable indépendante étant l'indicateur composite ESCS (X).

Deux séries de variables de la base de données du PISA sont donc mobilisées : d'une part, les variables attenantes aux scores en mathématiques (voir supra) et d'autre part, l'indicateur ESCS. Cet indicateur composite créé par l'OCDE prend en compte le niveau d'éducation du parent le plus instruit, la profession du parent la plus élevée selon les standards de *l'International Standard Classification of Education*⁴⁴ (ISCED), le niveau de vie de la famille et le montant des biens culturels disponibles au foyer⁴⁵ (OCDE, 2009). Il permet donc une mesure du niveau socioéconomique des élèves.

⁴⁴ <http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf>

⁴⁵ L'OCDE (*technical manual*) précise : *The ESCS scores were obtained as component scores for the first principal component with zero being the score of an average OECD student and one being the standard deviation across equally weighted OECD countries.* (OCDE, 2009, p. 16)

Inégalité des acquis de base

L'inégalité des acquis de base réfère à l'inégale accession à un niveau commun d'acquisition (Crahay, 2013). Il engage la fixation d'un seuil de référence de compétence minimale qui soit commun à chaque système éducatif étudié et donc, d'un référent international commun qui, bien que soumis à caution, est au cœur du projet PISA. Pour définir ce seuil, nous nous basons sur le niveau de compétences 2 du PISA, considéré par l'OCDE comme le niveau élémentaire de compétence :

« Au niveau 2, les élèves peuvent interpréter et reconnaître des situations dans des contextes qui leur demandent tout au plus d'établir des inférences directes. Ils ne peuvent puiser des informations pertinentes que dans une seule source d'information et n'utiliser qu'un seul mode de représentation. Ils sont capables d'utiliser des algorithmes, des formules, des procédures ou des conventions élémentaires pour résoudre des problèmes comprenant des nombres entiers. Ils peuvent interpréter les résultats de manière littérale » (OCDE, 2014 p. 75)

Il représente pour l'OCDE un seuil critique en deçà duquel les élèves risquent « d'être fortement pénalisés lorsqu'ils tentent de poursuivre des études supérieures et d'entrer dans la vie active par la suite » (OCDE, 2014, p. 75).

L'indicateur d'inégalité des acquis de base réfère donc, dans notre étude, à la proportion d'élèves sous le niveau de compétence 2 du PISA dans chaque système d'éducation⁴⁶. Il est aussi mobilisé dans nombre d'études (Cosnefroy & Rocher, 2004; Crahay, 2004; Demeuse & Baye, 2009; Mons, 2009; Rémond, 2006).

⁴⁶ Puisque nous prenons ici en compte la proportion d'élèves n'ayant pas acquis le niveau 2 de compétences PISA, les élèves ayant atteint le niveau de compétence 2 sont exclus..

3.4.2.2.3 Les variables de contrôle

Puisque l'objectif de ce second volet de recherche est d'analyser l'effet net des SABP sur l'efficacité et l'équité scolaires, nous mobilisons un ensemble de variables qui seront utilisées comme contrôle dans les analyses de régressions multiples. Sur base de la littérature, nous avons sélectionné trois variables de contrôle : 1) la part des dépenses publiques en éducation, 2) la variation du niveau socioéconomique des élèves au sein d'un même système éducatif et 3) la stratification des systèmes d'éducation. La littérature a, en effet, mis en exergue le rôle majeur joué par ces facteurs économiques, liés à la structure sociale des systèmes éducatifs (Mons, 2009) et à l'organisation de ces derniers, sur les degrés d'efficacité et d'équité scolaires.

1) La part des dépenses publiques en éducation rend compte de l'effort budgétaire des pays (des provinces ou des régions) en matière d'éducation (Duru-Bellat, Mons, & Suchaut, 2004; Hallak, 1996; Mons, 2004). Cet indicateur mobilisé dans plusieurs études, dont les études états-uniennes discutées au chapitre précédent (sous la forme des dépenses par élève), est appréhendé dans notre recherche en lien avec la richesse des pays (Produit Intérieur Brut - PIB) qui varie par ailleurs entre les pays, les provinces et les régions de notre échantillon. Il s'agit donc d'un indicateur relatif qui réfère à la part des dépenses publiques investie en éducation en proportion du PIB. Il a été obtenu via les bases de données et de la documentation internationales (voir supra).

2) Le milieu socioéconomique des élèves et son influence sur les performances scolaires a fait l'objet de long développement dans la littérature (Coleman et al., 1966; Jencks, 1972; Marks, 2005; OCDE, 2001, 2007, 2013a; Willms, 2004). Il est étudié de longue date comme un facteur d'inégalités au sein des systèmes d'éducation. Il est intégré dans notre recherche comme un indicateur de la structure sociale des systèmes éducatifs (Mons, 2004). Il se rapporte à la variance de l'indicateur ESCS disponible dans la base de données du PISA (voir supra) au sein des systèmes d'éducation :

« Si les pays laissent exprimer plus ou moins d'inégalités sociales de performances entre élèves, c'est avant tout en raison de deux facteurs d'importance voisine, qui sont assez peu corrélés, et viennent donc s'ajouter : l'existence d'inégalités interindividuelles en leur sein d'une part, le caractère plus ou moins différencié de leur système éducatif d'autre part » (Duru-Bellat, Mons & Suchaut, 2004, p. 129).

Même si le milieu socioéconomique des élèves est déjà pris en compte dans le calcul de l'indicateur d'inégalité des chances de résultats, il nous semble pertinent de le mobiliser en lien

avec les autres indicateurs d'équité. Cela nous permettra en effet d'observer si les variables d'*accountability* jouent sur des dimensions de l'équité et permettent d'atténuer le poids de l'origine socioéconomique sur les proportions d'élèves faibles ainsi que sur le différentiel de performance entre les élèves les plus performants et les élèves les plus faibles.

3) La citation de Duru-Bellat, Mons et Suchaut (2004) fait le pont avec la dernière variable de contrôle mobilisée. Il s'agit de l'indicateur de stratification horizontale et verticale, autrement nommé caractère différenciateur des systèmes éducatifs. Il est associé négativement à l'efficacité et l'équité des systèmes scolaires dans la plupart des études qui le prennent en compte (Duru-Bellat, Mons & Suchaut, 2004 ; Dupriez & Vandenberghe, 2004 ; Dupriez & Dumay, 2005; Le donné, 2017). Cet indicateur prend la forme, dans notre recherche, d'une variable agrégée construite à partir de trois informations⁴⁷ :

- a) le nombre de filières scolaires d'orientation rencontrées à l'âge de 15 ans. Ces filières étanches orientent les élèves vers des parcours différenciés. Cette information est fournie par l'OCDE (2013);
- b) deuxièmement, nous prenons en compte la pratique du redoublement, soit la proportion d'élèves ayant redoublé au moins une fois à 15 ans. Cette information est fournie par la base de données du PISA ;
- c) finalement, nous calculons la part de variance des scores des élèves expliquée par l'appartenance à un établissement scolaire. Statistiquement, cette mesure dispersive est appréhendée par le calcul du coefficient de variation intra-classe. Il s'agit du rapport entre la variance des scores des élèves au sein des établissements et la variance des scores des élèves entre établissements scolaires (variance intra/variance inter) (OCDE, 2013a). Il est calculé à partir des informations disponibles dans la base de données PISA 2012.

En guise de synthèse, nous reprenons dans le tableau ci-dessous (Tableau 8) l'ensemble des variables mobilisées dans les analyses réalisées que nous présentons dans les sections suivantes.

⁴⁷ Ces trois variables sont centrées réduites, puis le calcul de la moyenne est appliqué.

Tableau 8 : Variables mobilisées dans les analyses de la thèse

Variables dépendantes	Variables indépendantes	Variables de contrôle
(In)Efficacité scolaire	Composantes instrumentales de l' <i>accountability</i>	Dépenses publiques en éducation (en proportion du PIB des pays et des régions)
(In)Égalité de résultats	Types empiriques de SABP	Variation du milieu socioéconomique des élèves
(In)Égalités des chances de résultats		Stratification horizontale et verticale des systèmes scolaires
(In)Égalités des acquis de base		

3.4.2.4 Les méthodes d'analyse des données

Rappelons ici que l'objectif central de ce deuxième volet de recherche est d'analyser les effets des SABP sur l'efficacité et l'équité scolaires. Il vise la production d'évidences empiriques permettant de répondre à notre question centrale de recherche : quels sont les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation ?

Puisque nous avons créé deux séries de variables permettant de caractériser les SABP dans les systèmes éducatifs de notre échantillon, nous analysons, dans ce volet de recherche, les effets 1) des logiques instrumentales des SABP sur l'efficacité et l'équité scolaires (inégalités de résultats, des chances de résultats, des acquis de base) des systèmes d'éducation ; 2) des types empiriques sur ces dimensions.

D'abord, des analyses préliminaires sont réalisées (corrélation, régression simple, analyses de variance). Par la suite, des analyses multivariées (analyses de régressions multiples) devraient nous permettre d'analyser les effets nets des SABP sur les degrés d'efficacité et d'équité scolaires. Nous les décrivons dans les sous-sections suivantes en précisant les variables mobilisées, leur statut (variables dépendantes, variables indépendantes), les stratégies d'analyse de données et les tests réalisés aux fins de contrôle, le cas échéant. Nous terminons cette sous-section par quelques considérations méthodologiques sur les analyses réalisées.

3.4.2.4.1 Analyses des effets des composantes instrumentales des SABP sur l'efficacité et l'équité scolaires

Nous réalisons une série d'analyses bi-variées (corrélations de Pearson et régression simple) et multivariées (régressions multiples) permettant d'analyser tout d'abord l'effet des composantes instrumentales des SABP sur les degrés d'efficacité scolaire.

Analyses de corrélation et de régression simple

Pour analyser l'effet brut des composantes instrumentales des SABP⁴⁸ sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon, nous réalisons tout d'abord une série d'analyses bi-variées.

Des corrélations de Pearson permettent tout d'abord d'analyser le degré d'intensité de la relation entre composantes instrumentales des SABP, efficacité et équité scolaires (Haccoun & Cousineau, 2007; Howell, 2006).

Nous analysons par ailleurs les liens entre les variables de contrôle mobilisées dans les analyses de régressions multiples, et l'efficacité et l'équité scolaires, mais aussi entre composantes des SABP et nos variables de contrôle.

Finalement, les régressions linéaires simples visent à mettre en exergue la part de variabilité des indicateurs d'efficacité et d'équité scolaires (Y) expliquée par les composantes instrumentales des SABP (X). Il s'agit ici d'observer l'effet brut de ces dernières SABP sur le degré d'efficacité et d'équité scolaires (Bressoux, 2010; Howell, 2008; Haccoun & Cousineau, 2010).

Analyses de régressions multiples

Nous réalisons par ailleurs une série d'analyses de régressions multiples où l'effet net de chacune des composantes instrumentales des SABP (X) sur les degrés d'efficacité et d'équité scolaires (Y) est analysé de manière indépendante.

L'analyse des effets nets implique donc d'adopter un raisonnement « toutes choses égales par ailleurs » (Bressoux, 2010; Mons, 2017). Autrement dit, nous intégrons à l'analyse les trois facteurs présentés en amont dont il a été démontré (ou suggéré) qu'ils influencent les variables dépendantes : la part des dépenses publiques en éducation, la variation du niveau socioéconomique des élèves et la stratification des systèmes scolaires⁴⁹. Nous pouvons ainsi

⁴⁸ Pour rappel, cette première série de variables prend la forme de scores factoriels attribués à chaque système éducatif. Ils rendent compte de l'ampleur de chaque composante extraite des ACP au sein des systèmes d'éducation de l'échantillon (n=59).

⁴⁹ Du fait de la taille de notre échantillon, nous sommes limitée dans le nombre total de variables que nous pouvons intégrer aux modèles de régressions multiples (pour respecter les critères de ratio/cas variables observés dans la littérature, critère de dix cas pour une variable qui, bien que discutée, reste tolérée [Field, 2009]).

observer si les effets des composantes des SABP, le cas échéant, se maintiennent quand d'autres facteurs centraux sont pris en compte (Ho, 2008; Howell, 2008; Field, 2009; Tabachnick & Fidell, 2013).

Pour ces analyses de régressions multiples, nous utilisons la méthode de régression par bloc. Les composantes instrumentales des SABP sont d'abord intégrées (individuellement) au modèle, puis les variables de contrôle. Cette méthode nous permet d'analyser la variation des coefficients (coefficient de régression standardisé β) des composantes et la variation du coefficient de détermination du modèle (R^2) après ajout des variables de contrôle.

Précisons ici qu'une série de tests pour vérifier le respect des postulats des analyses de régression comme les tests de Kolmogorov-Smirnov et Shapiro-Wilk ont mis en exergue une forte asymétrie dans la distribution des variables dépendantes, plus particulièrement pour les indicateurs d'égalité de résultats (Howell, 2008; Field, 2009). Nous avons donc procédé à une transformation logarithmique des données pour l'ensemble des indicateurs d'efficacité et d'équité, puisque les autres transformations (racine carrée par exemple) se sont avérées insuffisantes pour pallier les problèmes de normalité des distributions. En outre, à la lumière des caractéristiques de nos données, la transformation log linéaire ne souffre pas de contre-indication majeure (Box & Cox, 1964; Howell, 2006).

3.4.2.4.2 Analyses des effets des types empiriques de SABP sur l'efficacité et l'équité scolaires

Nous analysons dans un second temps l'effet des types de SABP sur les degrés d'efficacité et d'équité scolaires. Cette série d'analyses nous permet donc de dépasser une analyse des composantes instrumentales en termes indépendants. Nous observons ainsi dans quelle mesure les types de SABP, au sein desquels ces dernières sont articulées selon des configurations différenciées en fonction des types, ont un impact sur l'efficacité et l'équité scolaires des systèmes éducatifs de notre échantillon.

La même logique d'analyse des effets nets, puis bruts, est appliquée pour les types empiriques de SABP. Nous réalisons tout d'abord des analyses de variance (ANOVA) puis des analyses de régression multiple.

Analyses de variance à un critère de classification (ANOVA)

Les analyses de variance à un facteur de classification (ANOVA) mobilisent les types empiriques d'*accountability* (échelle nominale à plusieurs modalités) comme facteur de classification, les indicateurs d'efficacité et d'équité scolaires comme variables dépendantes (Y). Ici, l'analyse porte sur les moyennes des types empiriques. On cherche donc à savoir s'il existe des différences statistiquement significatives en ce qui a trait au degré moyen d'efficacité et d'équité scolaires des types empiriques de SABP construits (Howell, 2008).

Des tests de Levene et Kruskal-Wallis ont été réalisés pour vérifier le postulat d'égalité des variances (ANOVA) (Ho, 2008; Howell, 2008; Field, 2009; Tabachnick & Fidell, 2013). Par ailleurs, les différences entre les moyennes des types empiriques mises en exergue par l'ANOVA, en d'autres termes, les contrastes entre types empiriques, ont été testés au moyen des tests de Tukey et de Bonferroni. Nous rapportons les résultats des tests de Tukey recommandé pour des petits échantillons (Howell, 2008).

Analyses de régressions multiples

Pour les analyses de régressions multiples mobilisant les types de SABP comme variable indépendante (X), la même stratégie d'analyse par bloc est utilisée.

Les types de SABP sont d'abord intégrés au modèle, puis les variables de contrôle. Ces analyses impliquent toutefois un raisonnement différent que lorsque la variable indépendante est de type continu. L'intégration d'une variable de type nominale (à plusieurs modalités) comme variable indépendante (ou critère) implique de choisir une classe de référence. Les effets des autres types sont donc interprétés en référence à cette classe (Field, 2009).

Nous choisissons ici comme modalité de référence le type de SABP où l'*accountability* est la plus faible. Il est donc exclu des analyses et l'interprétation des effets constatés, le cas échéant, va comme suit : les effets des types empiriques analysés diffèrent-ils de ceux de la classe de référence ?

Finalement, en termes d'interprétation des résultats des analyses de régressions multiples, si nous observons les tailles d'effet des modèles, nous nous intéresserons surtout aux coefficients standardisés des variables introduites et à leur caractère significatif. Il s'agit d'analyser si oui ou non les variables d'*accountability* (composantes instrumentales des SABP et types empiriques) ont un impact sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs, et si cet effet se maintient toute chose étant égale par ailleurs. De fait, nous ne raisonnons pas en termes de prédiction. Notre position est explicitée par Howell (2008) :

Dans la plupart des applications de la régression (...) faire une prédiction spécifique ne nous intéresse pas particulièrement. Nous pourrions certes trouver intéressant de connaître la relation entre le revenu familial et les performances scolaires, mais il est peu probable que nous utilisions la mesure du revenu familial d'un enfant déterminé pour prédire ses performances scolaires. Nous nous intéressons généralement bien davantage aux principes généraux qu'aux prédictions individuelles. Une équation de régression peut nous donner des informations utiles sur ces principes généraux, même si n'allons peut-être jamais l'utiliser réellement pour prédire un cas spécifique (Howell, 2008, p. 248-249).

Avant de discuter quelques considérations méthodologiques, nous synthétisons les analyses réalisées dans ce deuxième volet de recherche empirique dans le Tableau 9.

Tableau 9 : Variables intégrées aux modèles de régressions multiples

Méthodes d'analyse	Variables dépendantes (Y)	Variables indépendantes (X)/facteur de classification	Variables de contrôle (le cas échéant)	Tests complémentaires
<i>1^{re} série d'analyses — effets de composantes instrumentales sur l'efficacité et l'équité scolaires</i>				
1) Corrélations de Pearson	Indicateurs d'efficacité (échelle d'intervalle)	Scores factoriels (composantes des SABP – échelle d'intervalle)		Kolmogorov-Smirnov Shapiro-Wilk
2) Analyses de régression simple	Indicateurs d'équité (variables continues) : 1) égalité de résultats 2) égalité des chances de résultats 3) égalité des acquis de base			
3) Analyses de régressions multiples	Indicateurs d'efficacité (échelle d'intervalle)	Scores factoriels (composantes des SABP – échelle d'intervalle)	1) Part des dépenses publiques en éducation 2) Structure sociale des systèmes éducatifs 3) Stratification horizontale et verticale	
	Indicateurs d'équité (variables continues) : 1) égalité de résultats 2) égalité des chances de résultats 3) égalité des acquis de base		IDEM (**lorsque l'indicateur d'égalité des chances de résultats est intégré aux analyses, la structure sociale des systèmes éducatifs n'est pas ajoutée comme contrôle)	
<i>2^e série d'analyses : effets des types de SABP sur l'efficacité et l'équité scolaires</i>				
1) ANOVA (1 facteur de classification)	Indicateurs d'efficacité (échelle d'intervalle)	Types empiriques de SABP (échelle nominale à plusieurs modalités)		Levene Kruskal-Wallis Bonferroni Tukey
	Indicateurs d'équité (échelle d'intervalle) : 1) égalité de résultats 2) égalité des chances de résultats 3) égalité des acquis de base			
2) Analyses de régressions multiples	Indicateurs d'efficacité (échelle d'intervalle)	Types empiriques de SABP (variable nominale – plusieurs modalités, dont une modalité de référence)	1) Part des dépenses publiques en éducation 2) Structure sociale des systèmes éducatifs 3) Stratification horizontale et verticale	

	Indicateurs d'équité (échelle d'intervalle) 1) égalité de résultats 2) égalité des chances de résultats (échelle d'intervalle) 3) égalités des acquis de base	IDEM	IDEM (**lorsque l'indicateur d'égalité des chances de résultats est intégré aux analyses, la structure sociale des systèmes éducatifs n'est pas ajoutée comme contrôle)	
--	--	------	---	--

3.4.2.5 Considérations méthodologiques

Nous terminons ce chapitre sur quelques considérations méthodologiques en lien avec la taille de notre échantillon et avec l'usage de la méthode des moindres carrés ordinaires (MCO) comme base des analyses réalisées.

Premièrement, comme nous l'avons souligné, la taille de notre échantillon peut apparaître relativement faible bien qu'acceptable. Si nous ne sommes pas la seule à réaliser ce type d'études avec un échantillon similaire (voir Mons, 2004; Dupriez & Dumay, 2011; etc.), quelques précisions s'imposent. La taille de l'échantillon joue particulièrement sur deux aspects : les tailles d'effets des modèles et les erreurs de type 1.

Comme mis en exergue par nombres d'auteurs (Ho, 2008; Howell, 2008; Bressoux, 2008; Field, 2009; Tabachnick & Fidell, 2013), la taille de l'échantillon a un impact sur la puissance des tests réalisés et sur les probabilités associées aux erreurs de type 1. En effet, un petit échantillon limite fortement la puissance des tests et influe sur la significativité des coefficients de détermination des modèles :

“The size of the sample has a direct impact on the statistical power of the significance testing in multiple regression. Power in multiple regression refers to the probability of detecting as statistically significant a specific level of R square, or a regression coefficient at a specified significance level and a specific sample size (Hair et al., 1995). Thus, for a desired level of power and with a specified number of independent variables, a certain sample size will be required to detect a significant R-square at a specified significance level (see Cohen and Cohen, 1983, for sample size calculations)” (Ho, 2006, p. 247)

De fait, un petit échantillon augmente la probabilité d'erreurs de type 1 et altère le seuil de significativité des tests d'hypothèses. Autrement dit, il y a davantage de probabilité de rejeter l'hypothèse d'un effet des variables indépendantes sur les variables dépendantes, alors que cette hypothèse est vraie. Néanmoins, comme précisé par Field (2009), les règles en matière de taille d'échantillon sont relativement floues, qui plus est, cette question renvoie à des réalités différenciées lorsqu'il est question de travailler à l'échelle des systèmes éducatifs. Il nous semble donc que l'important est de garder à l'esprit les implications de travailler à partir d'un petit échantillon.

“Well, how much is enough? You’ll find a lot of rules of thumb floating about, the two most common being that you should have 10 cases of data for each predictor in the model, or 15 cases of data per predictor. So, with five predictors, you’d need 50 or 75 cases respectively (depending on the rule you use). These rules are very pervasive (even I used the 15 cases per predictor rule in the first edition of this book) but they over simplify the issue considerably. In fact, the sample size required will depend on the size of effect that we’re trying to detect (i.e. how strong the relationship is that we’re trying to measure) and how much power we want to detect these effects. The simplest rule of thumb is that the bigger the sample size, the better!” (Field, 2009, p. 223)

Finalemment, rappelons que les analyses réalisées dans le second volet de recherche reposent sur la méthode des MCO. On postule donc l’existence d’une relation linéaire entre ces variables. Toutefois, nous pouvons supposer d’une part que les relations entre SABP, efficacité et équité scolaires pourraient être soumises à des effets de seuils et, d’autre part, que la relation linéaire entre dimensions reste sensiblement faible. Si la question des effets de seuils n’est pas analysée dans notre thèse, ces questions seront abordées plus avant dans la discussion de nos résultats et des limites de la thèse.

3.5 Conclusion

Ce chapitre avait pour visée d’explicitier la démarche méthodologique déployée dans notre recherche doctorale pour répondre à la question de recherche centrale des effets des SABP sur l’efficacité et d’équité scolaires. Après avoir identifié un manque dans la littérature, en ce qui a trait aux études quantitatives comparatives inter-systèmes éducatifs portant sur cette question dans un contexte hors États-Unis, nous avons mis en exergue l’effort nécessaire de conceptualisation de la notion d’*accountability* et la pertinence d’une entrée analytique par l’instrumentation.

En découlent les objectifs de notre recherche quantitative comparative internationale qui vise d’une part à caractériser les SABP dans les systèmes d’éducation d’Europe et du Canada et d’autre part à analyser leurs effets sur leurs degrés d’efficacité et d’équité scolaires. L’idée centrale est donc d’analyser les instruments au cœur des SABP en action et, dès lors, leurs effets en termes d’efficacité et d’équité scolaires.

Nous déployons donc une démarche méthodologique reposant principalement sur une analyse secondaire des données du PISA 2012, base de données comparatives parmi les plus conséquentes en éducation.

Notre chantier empirique repose ainsi sur deux volets distincts, mais interreliés. Le premier présente un intérêt descriptif et analytique certain. Axé sur l'opérationnalisation de la notion de SABP, il permet d'une part d'analyser les configurations d'instruments au cœur des SABP, et d'autre part de classer les systèmes éducatifs d'Europe et du Canada sur cette base et de construire ainsi une typologie empirique originale des SABP. D'un point de vue opérationnel, il aboutit à la construction de variables inédites mobilisées dans un second volet de recherche pour analyser les effets des SABP.

Notre deuxième volet analytique ne se contente donc pas d'interroger les effets des SABP en termes d'efficacité, mais questionne aussi leur impact sur différentes figures des inégalités scolaires. Ici, nous interrogeons les effets bruts des SABP, mais aussi leurs effets nets en introduisant différentes variables de contrôle : économique, liée à la structure sociale et organisationnelle des systèmes éducatifs.

DEUXIÈME PARTIE : RÉSULTATS DE RECHERCHE

CHAPITRE 4 : LA DIVERSITÉ DES SYSTÈMES D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE EN EUROPE ET AU CANADA – UNE APPROCHE EMPIRIQUE PAR LES INSTRUMENTS/MODES DE RÉGULATION

Ce chapitre présente les résultats du premier volet empirique de recherche. Comme précisé dans notre cadrage méthodologique, ce premier volet de recherche constitue un préalable nécessaire à l'analyse des effets des SABP sur les degrés d'efficacité et d'équité scolaires et poursuit plusieurs objectifs. Il vise à 1) caractériser les systèmes d'*accountability* basés sur la performance (SABP) dans les systèmes éducatifs d'Europe et du Canada ; 2) classifier ces derniers sur base des modes de régulation mis en œuvre au sein des établissements scolaires ; et 3) créer des variables permettant d'analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires.

Il repose sur les postulats que nous avons formulés précédemment. Nous posons en effet que 1) les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes et 2) permettent de différencier des types empiriques de SABP.

En cohérence avec ces postulats, nous avons dès lors adopté une démarche visant à prendre en compte les instruments mobilisés au sein des établissements scolaires pour caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada. Deux types d'analyses, des analyses en composantes principales (ACP) et des analyses de classification hiérarchique (ACH), ont donc été réalisées sur les variables clefs de la base de données du PISA 2012. Les premières (ACP) nous ont permis d'extraire trois composantes instrumentales des SABP (une composante d'évaluation et de régulation selon la performance, d'incitatifs pour les enseignants, de compétition entre écoles et d'amélioration de leurs performances), les secondes (ACH), de construire des types empiriques d'instrumentation des SABP. Chacune, finalement, a permis la construction des variables originales permettant, dans notre deuxième volet de recherche, d'analyser les effets des SABP (composantes instrumentales et types) sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon.

Le Tableau 10 ci-dessous reprend les principaux éléments au cœur de ce premier volet de recherche.

Tableau 10 : Tableau synoptique du volet 1 de recherche

	Volet empirique 1	
Questionnement central de la thèse	Quels sont les effets des SABP sur le degré d'efficacité et d'équité scolaires des systèmes d'éducation ?	
Sous-question de recherche	Quelles formes revêtent les SABP dans les systèmes d'éducation d'Europe et du Canada ?	
Objectifs	<u>Descriptif et analytique</u> : 1) caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada à partir de la gamme et de l'étendue des modes de régulation mis en œuvre au sein des établissements scolaires; 2) classifier les systèmes d'éducation de l'échantillon sur cette base <u>Opérationnel</u> : 3) construction de variables permettant d'analyser les effets des SABP sur l'efficacité et l'équité scolaires.	
Postulats	1) Les instruments au cœur des SABP sous-tendent des logiques de régulation distinctes. 2) Les instruments permettent de différencier des types empiriques de SABP.	
Concepts clefs	Modes de régulation institutionnelle, SABP	
Méthodes d'analyse des données	Analyses en composantes principales [ACP]	Analyses de classification hiérarchique [<i>Cluster</i>]
Variables mobilisées	1) Dispositifs et instruments aux fins d'assurance qualité et d'amélioration des écoles (SC39) 2) Mécanismes de reddition de comptes (SC19) 3) Usage des résultats des élèves par les établissements scolaires (SC18) 4) Suites, conséquences des évaluations des enseignants (SC31)	Composantes instrumentales (scores factoriels attribués aux systèmes d'éducation de l'échantillon) créées via les ACP réalisées sur les variables présentées ci-contre.
Variables créées	Composantes instrumentales des SABP	Types d'instrumentation des SABP selon les systèmes éducatifs
Types de variable créée	Variable continue — échelle d'intervalle (scores factoriels)	Variable nominale — échelle nominale à plusieurs modalités

Dans une première section, nous présentons tout d'abord les résultats des analyses en composantes principales. La seconde section présente les résultats des analyses de classification hiérarchique.

SECTION 1 : Composantes instrumentales des systèmes d'*accountability* basés sur la performance

Cette section présente les résultats des analyses en composantes principales (ACP). Dans les sous-sections qui suivent, après avoir analysé les paramètres centraux des ACP, nous décrivons les composantes extraites et réalisons un descriptif comparatif des systèmes éducatifs d'Europe et du Canada sur cette base. Nous proposons finalement une interprétation des composantes extraites de nos analyses.

4.1 Résultats des analyses en composantes principales

Pour rappel, les variables incluses à l'analyse sont au nombre de 17 (Tableau 11). Elles sont regroupées dans la base de données du PISA selon quatre séries ayant trait : 1) aux dispositifs et aux instruments mobilisés par les établissements scolaires aux fins d'assurance qualité et d'amélioration des écoles 2) au suivi des résultats des écoles (suivi par une autorité administrative et publication des résultats) 3) à l'usage des résultats des élèves par les écoles 4) aux incitatifs ou aux conséquences qui font suite aux évaluations des enseignants.

Tableau 11 : Rappel, variables d'intérêt, instruments des SABP

VARIABLES DU PISA 2012 —INSTRUMENTS DES SABP			
Série 1	Série 2	Série 3	Série 4
Dispositifs ou instruments mobilisés dans les établissements aux fins d'assurance qualité et d'amélioration des écoles	Suivi des performances des écoles et publication des résultats des élèves	Usage des résultats des élèves au niveau des établissements scolaires	Suites, conséquences (incitatifs) des évaluations des enseignants
<i>Which of the following measures aimed at quality assurance and improvement do you have in your school?</i> (SC 39)	<i>In your school, are achievement data used in any of the following <accountability procedures</i> (SC 19)	<i>In your school, are assessments of students in <national modal grade for 15-year-olds> used for any of the following purposes?</i> (SC 18)	<i>To what extent have appraisals of and/or feedback to teachers directly led to the following?</i> (SC 31)
Consultation d'experts dans les six derniers mois (Q09 ⁵⁰)	Suivi par une autorité administrative (Q02)	Comparaison des performances de l'école à celles des autres écoles (Q08)	Attribution de responsabilités professionnelles (Q06)
Évaluation externe (Q06)	Publication des résultats (Q01)	Évaluation de la performance des enseignants (Q06)	Reconnaissance publique de la part du chef d'établissement (Q05)
Évaluation interne (Q05)		<i>Monitoring</i> des performances des écoles d'une année sur l'autre (Q05)	Implications en termes d'avancement de carrière (Q04)
Enregistrement systématique des données incluant les taux de présence et de graduation des enseignants et des élèves, résultats aux tests, participation des enseignants à des activités de développement professionnel (Q03)		Comparaison des performances de l'école aux performances régionales ou nationales (Q04)	Activités de développement professionnel (Q03)
Spécification des standards de performance à atteindre (Q02)			Bonus financier (Q02)
			Changement de salaire (Q01)

⁵⁰ Cette information renvoie à l'identification des variables pour chaque série (SC39, SC19, SC31, SC18) dans la base de données PISA « écoles » 2012 tel que précisé dans le « codebook » qui accompagne cette dernière.

L'examen des paramètres centraux et des résultats des ACP réalisées sur ces variables atteste que la solution factorielle est de qualité fort acceptable (Tableau 12). Les résultats du test de sphéricité de Bartlett ($\chi^2 (136) = 798,93, p = 0,000$) indiquent la présence de corrélations inter-items. Nos données forment donc un ensemble cohérent (Y., Pras, & Roux, 1993) permettant de chercher des dimensions communes qui y aient un sens (Houdayer & Fontane, 2001). Ce test étant très sensible au ratio variables/cas, nos résultats sont d'autant plus encourageants. La mesure de Kaiser-Meyer-Olkin atteste en outre de l'adéquation de la solution factorielle (KMO = 0,852).

Tableau 12 : Résultats de l'ACP, indice de Kaiser-Meyer-Olkin et test de Bartlett

Indice KMO et test de Bartlett		
Indice de Kaiser-Meyer-Olkin pour la mesure de la qualité d'échantillonnage.		0,852
Test de sphéricité de Bartlett	Khi-deux approx.	798,973
	ddl	136
	Signification	0,000

Le critère de Kaiser (valeurs propres supérieures à 1,0) suggère de retenir trois composantes qui permettent d'extraire 70,97 % de l'inertie totale⁵¹. La proportion de variance totale extraite sur ces trois composantes n'engage donc pas une perte trop importante d'information, qui s'élève ici à moins de 30 %. Ces trois composantes principales expliquent respectivement 48,74 %, 13,26 % et 8,97 % de l'inertie.

Elles renvoient à des composantes instrumentales des SABP construites par des variables distinctes, que nous présentons ci-après : 1) une composante d'évaluation selon la performance; 2) une composante se rapportant aux incitatifs et enjeux articulés aux évaluations des enseignants; et 3) une composante axée sur la comparaison des résultats des établissements scolaires et dispositifs internes d'amélioration de leurs performances.

⁵¹ L'inertie est égale à la somme des variances de variables étudiées.

4.1.1 Première composante instrumentale : évaluation et régulation selon la performance

La première composante explique la plus forte proportion de variance, elle est construite par sept variables (Tableau 13). Y sont regroupées les variables ayant trait aux dispositifs d'assurance qualité mis en place au sein des établissements scolaires (à l'exception de la consultation d'experts externes), ainsi que celles qui se rapportent au suivi des résultats des élèves et des écoles par une autorité administrative à l'usage des résultats des élèves pour suivre leur progrès d'une année à l'autre et pour évaluer les performances des enseignants.

Les indices de saturation permettent de rendre compte de la contribution différenciée des variables à la construction des composantes. Nous observons que l'enregistrement des données de performance et de gestion par les établissements scolaires, ainsi que l'usage des résultats des élèves par ces derniers pour *monitorer* leurs performances, apportent les contributions les plus importantes (indices de saturation de 0,857 et 0,821), l'évaluation interne et la spécification des standards à atteindre, les contributions les plus faibles (indices de saturation de 0,626 et 0,672 respectivement). Le coefficient alpha de Cronbach est en outre élevé ($\alpha = 0,885$). Il permet d'attester de la fiabilité de la cohérence interne de la composante.

Tableau 13 : Indices de saturation des variables sur composante d'évaluation selon la performance

Item	Saturation	Alpha de Cronbach
39Q03	Enregistrement des données de performance	0,857
18Q05	Usage des résultats des élèves par les établissements pour suivre leurs progrès d'une année à l'autre	0,821
39Q06	Évaluation externe	0,743
19Q02	Suivi des résultats des élèves par une autorité hiérarchique	0,700
18Q06	Usage des résultats des élèves pour évaluer la performance des enseignants	0,695
39Q02	Spécification des standards de performances	0,672
39Q05	Évaluation interne	0,626

Cette composante instrumentale réfère à une dimension fondamentale des systèmes d'*accountability* basés sur la performance, à savoir l'évaluation et le suivi de la performance des acteurs. Les résultats des élèves constituent ici la pierre angulaire d'un ensemble d'instruments et de mécanismes mis en œuvre au niveau des établissements scolaires dans une perspective de pilotage par les résultats et d'amélioration des performances des acteurs (élèves, enseignants, établissements). Les mécanismes de suivi des résultats des élèves au niveau des écoles elles-mêmes, et par une autorité administrative, y sont centraux. Dans cette logique c'est donc le niveau de l'école qui est le niveau d'action, l'acteur ou l'entité centrale, ciblé dans une visée d'amélioration du système et d'augmentation des performances.

Précisons ici que si la variable ayant trait à l'usage des résultats des élèves pour évaluer les performances des enseignants participe clairement à la construction de cette composante, elle enregistre par ailleurs un indice de saturation relativement élevé (de 0,532) sur la deuxième composante extraite de nos analyses (voir infra). En cela, cette variable peut être considérée comme une variable complexe. Néanmoins, il n'est pas pertinent de la supprimer des analyses car elle rend compte d'un aspect fort intéressant pour notre propos, à savoir la prise en compte des résultats des élèves pour évaluer la performance des enseignants.

4.1.2 Deuxième composante instrumentale : enjeux et incitatifs directs pour les enseignants

La seconde composante extraite des ACP explique quant à elle 13,26 % de l'inertie totale. Elle est construite par cinq variables (Tableau 14) ayant trait au suivi et aux conséquences des évaluations du corps enseignant, en d'autres termes, aux enjeux et aux incitatifs directs pour les enseignants articulés à leurs évaluations. Le questionnaire PISA permet en effet de renseigner les conséquences matérielles et sociales qui sont directement attachées aux évaluations des enseignants⁵².

Ce sont ici les incitatifs financiers qui apportent la contribution la plus importante à la construction de cette composante (indice de saturation de 0,899 pour l'attribution de bonus financiers et de 0,831 pour le changement de salaire). À l'inverse, deux variables apportent les

⁵² Pour rappel, la formulation du questionnaire du PISA va comme suit: *“To what extent have appraisals of and/or feedback to teachers directly led to the following?”*

contributions les plus faibles : la reconnaissance publique de la part du chef d'établissement et les implications en termes d'avancement de carrière (indices de saturation respectifs : 0,645 et 0,648). La valeur de l'alpha de Cronbach est aussi élevée pour cette deuxième composante ($\alpha = 0,879$), ce qui indique une bonne cohérence interne de cette composante.

Tableau 14 : Indices de saturation des variables de la composante incitatifs et enjeux pour le corps enseignant

Item		Saturation	Alpha de Cronbach
31Q02	Bonus financier	0,899	0,879
31Q01	Changement de salaire	0,831	
31Q06	Attribution de responsabilités professionnelles censées rendre la profession plus attractive	0,724	
31Q05	Reconnaissance publique de la part du chef d'établissement	0,648	
31Q04	Avancement de carrière	0,645	

Cette composante renvoie ainsi clairement aux conséquences et aux enjeux qui sont articulés aux évaluations des enseignants. Ce sont les enseignants qui sont ici les acteurs, ou l'entité centrale visée, au cœur de mécanismes visant l'amélioration des performances organisationnelles. Par ailleurs, comme nous l'avons précisé dans la sous-section précédente, la variable se rapportant à l'usage des résultats des élèves pour évaluer les performances des enseignants présente une forte corrélation avec cette composante. Cela suggère que cette composante instrumentale, en d'autres termes, les incitatifs attachés à l'évaluation des enseignants, n'est pas déconnectée de l'enjeu de performance (des enseignants, des élèves), bien au contraire.

4.1.3 Troisième composante instrumentale : comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires

Finally, the third and last component extracted from the ACP is the one that explains the lowest proportion of total inertia (8,97 %). It is constructed from five variables (Tableau 15), three of which relate to comparison instruments and, indirectly, to competition between schools based on their performance: publication of school results, use of student results by schools to compare their performance with those of other schools, as well as national and regional performance. The other two variables deal, on the one hand, with the use of external experts for the improvement of schools and, on the other hand, with professional development opportunities offered to teachers as a result of their evaluation.

Here, the consultation of external experts for quality assurance and school improvement as well as the publication of results show the highest saturation indices (respectively 0,783 and 0,772); the use of student results by schools to compare their performance with those of other schools and professional development opportunities for teachers, the lowest (0,605 and 0,709).

It is also the instrumental component that presents the lowest statistical coherence. The Cronbach's alpha value is nevertheless superior to 0,70 ($\alpha = 0,767$) which, according to the work of Ho (2006) and Frisbie (1988), attests to a strong internal coherence.

Tableau 15 : Indices de saturation des variables de la composante de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires

Item		Saturation	Alpha de Cronbach
39Q09	Consultation d'experts — amélioration de l'école	0,783	0,767
19Q01	Publication des résultats	0,772	
18Q04	Usage des résultats des élèves par les établissements aux fins de comparaison aux performances nationales ou régionales	0,716	
31Q03	Suite des évaluations des enseignants — activités de développement professionnel	0,709	
18Q08	Usage des résultats des élèves par les établissements aux fins de comparaison aux performances des autres écoles	0,605	

Cette composante réfère, *a priori*, à des instruments distincts qui, pourtant, se combinent pour construire cette dimension. Cette logique, qui vise l'amélioration des performances des établissements scolaires, s'appuie en effet sur deux éléments centraux : d'une part, la mise en comparaison des établissements scolaires sur base de leurs performances, d'autre part, la mise en œuvre de dispositifs d'intervention au sein des écoles. Le niveau d'action ciblé, ou l'entité centrale visée, y est donc aussi, à l'instar de la première composante instrumentale, l'établissement scolaire.

Ces trois composantes extraites, nous avons souhaité observer leur ampleur au sein des systèmes éducatifs de notre échantillon. Les scores factoriels attribués aux systèmes éducatifs permettent, en effet, de quantifier l'ampleur des composantes instrumentales en leur sein. Ils nous permettent donc de comparer, à titre descriptif, les systèmes éducatifs d'Europe et du Canada sur cette base.

4.2 Une comparaison des systèmes éducatifs d'Europe et du Canada selon les différentes composantes instrumentales des systèmes d'*accountability* basés sur la performance

Nous présentons ici le comparatif des systèmes éducatifs de l'échantillon sur base des scores factoriels qui leurs sont attribués⁵³.

La Figure 4 présente ce comparatif descriptif pour la composante d'évaluation et de régulation en fonction de la performance. On observe que, situés dans le premier quartile de la distribution, quinze systèmes éducatifs, dont les systèmes éducatifs de la Grèce, les Länders allemands du Bade-Wurtemberg, de l'Hesse et du Bayern ainsi que la France, l'Italie, le Danemark, les communautés francophones et germanophones de Belgique ou encore l'Autriche se voient attribuer les scores factoriels aux valeurs les plus faibles. Cette composante instrumentale d'évaluation selon la performance est donc de plus faible ampleur au sein de ce groupe que dans le reste de l'échantillon.

Les systèmes d'éducation de la Colombie-Britannique, de la Finlande, de l'Irlande ou encore du Québec font, quant à eux, partie des quatorze systèmes, situés dans le deuxième quartile de la distribution, où cette composante instrumentale est d'intensité faible à moyenne. Les systèmes d'éducation de l'Alberta, de la Lituanie, de l'Angleterre se situent quant à eux dans le troisième quartile de la distribution. Dans ce groupe de seize systèmes d'éducation, cette composante est donc d'intensité moyenne à forte.

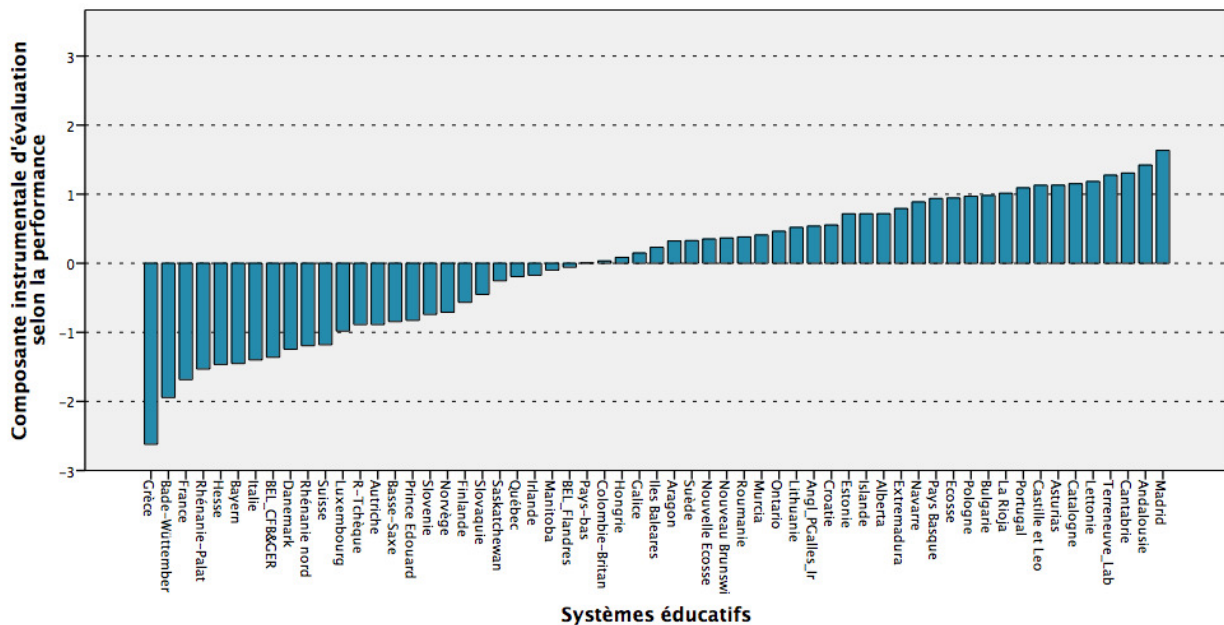
Enfin, d'après nos données, cette composante a la plus forte ampleur dans treize systèmes éducatifs, dont l'Écosse, la Pologne, la Bulgarie, les régions espagnoles de Madrid et de l'Andalousie ainsi que dans les provinces canadiennes de Terre-Neuve et Labrador.

Nous avons par ailleurs souhaité observer, à titre indicatif, si des différences statistiquement significatives pour cette composante instrumentale des SABP existent entre ces groupes de systèmes éducatifs classés par quartile de la distribution. Les résultats de l'ANOVA à un facteur

⁵³ Pour rappel, les scores factoriels prennent la forme de variables standardisées pour lesquelles la moyenne est de 0 l'écart-type de 1. Les systèmes éducatifs qui se voient attribuer les scores factoriels aux valeurs les plus élevées sont ceux où les composantes y sont les plus développées, ceux où ces valeurs sont les plus faibles sont ceux où les composantes instrumentales y sont les moins développées.

de classification⁵⁴ [$F(3,55) = 174,78; p = 0,00$] et du test de Tukey attestent en effet de la présence de différences statistiquement significatives entre chacun des quatre groupes.

Figure 4 : Composante instrumentale d'évaluation et de régulation selon la performance, comparaison des systèmes éducatifs de l'échantillon



Le comparatif des systèmes d'éducation pour la composante instrumentale ayant trait aux conséquences et aux suites de l'évaluation des enseignants est illustré par la Figure 5. Il indique que cette composante est de moindre ampleur dans quinze systèmes éducatifs, situés dans le premier quartile de la distribution : les provinces canadiennes de l'Ontario, de Terre-Neuve et Labrador, les régions d'Espagne de la Rioja et de la Galice, ainsi qu'en Belgique francophone et germanophone.

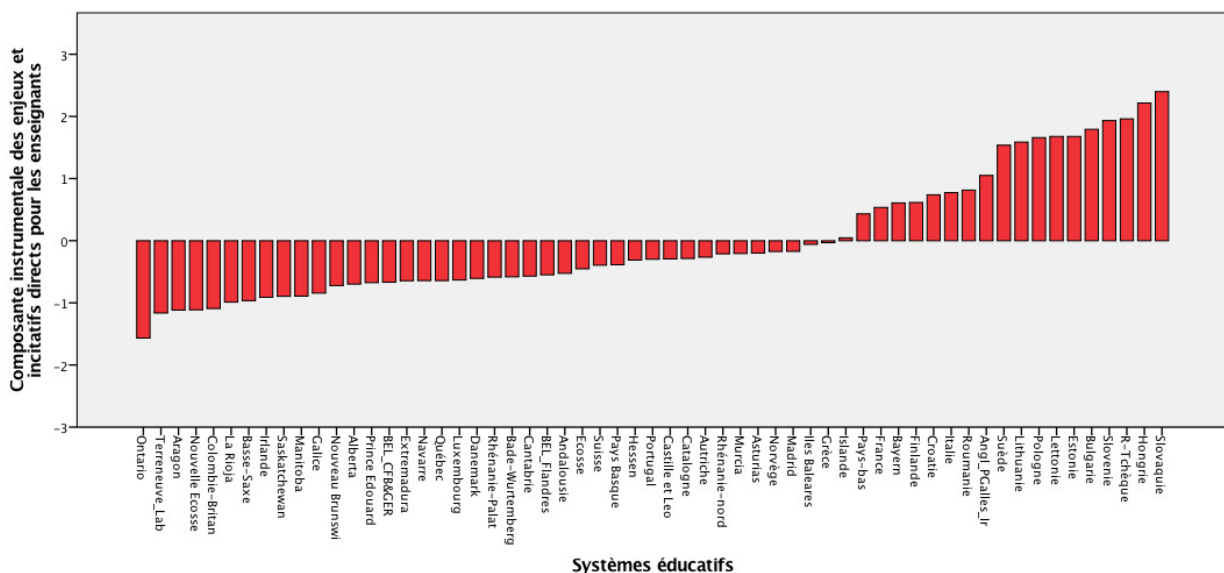
La composante des incitatifs pour les enseignants est par ailleurs d'intensité faible à moyenne dans les systèmes d'éducation du Québec, du Danemark, de l'Écosse ou encore de la Suisse, situés dans le deuxième quartile de la distribution ; d'intensité moyenne à forte au Portugal, en Autriche, aux Pays-Bas ou encore en France, situés dans le troisième quartile de la distribution.

⁵⁴ Une ANOVA à un facteur de classification a été réalisée pour observer la présence de différences statistiquement significatives entre les groupes des systèmes éducatifs classés par quartile de la distribution des scores factoriels (les quartiles ayant été créés automatiquement via le logiciel d'analyse). La variable indépendante (Y) est le score factoriel attribué à chaque système éducatif pour chaque composante instrumentale, la variable indépendante (X) les groupes de systèmes éducatifs. Les résultats des tests de Bonferroni et de Tukey attestent de la présence de différences statistiquement significatives entre les quatre groupes de systèmes éducatifs. La même opération a été réalisée pour chaque composante instrumentale des SABP.

Enfin, ce sont les systèmes éducatifs de la Suède, de l'Angleterre ainsi que des pays de l'est de l'Europe, tels que la Slovaquie, la Hongrie ou la République tchèque, qui enregistrent les valeurs les plus fortes.

Les résultats de l'ANOVA [$F(3,55) = 140,89 ; p = 0,00$] et du test de Tukey confirment par ailleurs la présence de différences statistiquement significatives entre ces quatre groupes.

Figure 5 : Composante instrumentale des enjeux et incitatifs directs pour les enseignants, comparaison des systèmes éducatifs de l'échantillon



Enfin, la Figure 6 illustre le comparatif pour la composante instrumentale de comparaison externe et dispositifs d'amélioration interne des performances de l'école.

Ce sont ici l'ensemble des régions d'Espagne qui se voient attribuer les scores factoriels aux valeurs les plus faibles, avec les systèmes éducatifs de la Finlande, de la Belgique francophone et germanophone et les *Lands* allemands de la Basse-Saxe et de la Rhénanie-Palatinat. Ils font partie, en effet, des quatorze systèmes éducatifs situés dans le premier quartile de la distribution.

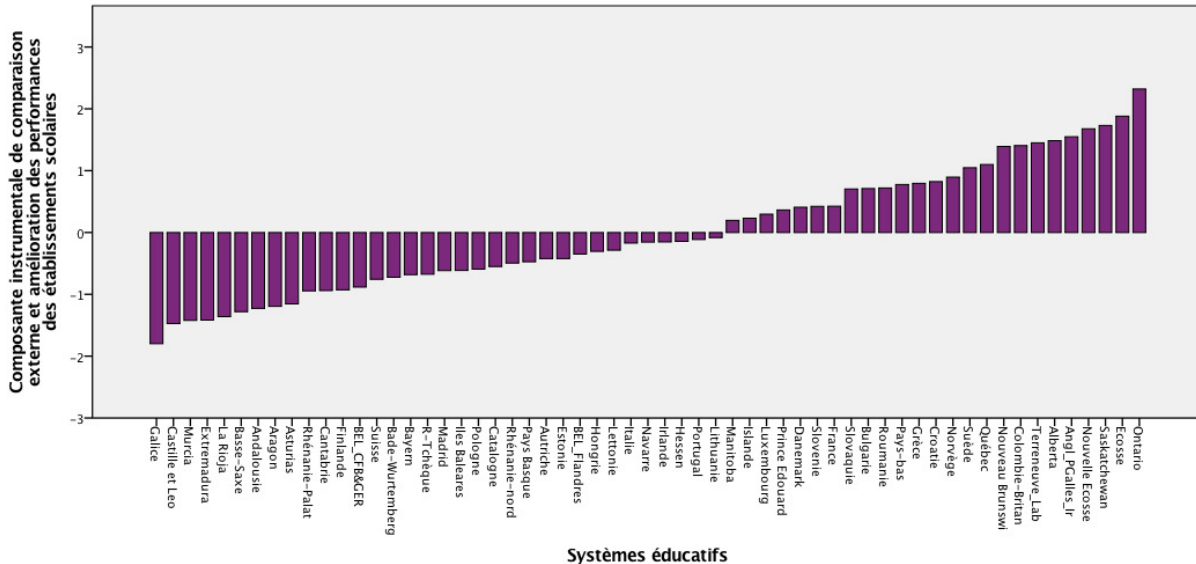
Les systèmes d'éducation de la communauté flamande de Belgique, de la Hongrie, de l'Italie ou encore de la région d'Espagne de Navarre se situent parmi les quinze systèmes d'éducation dans le deuxième quartile de la distribution. Cette composante est par ailleurs, d'après nos résultats, de plus forte ampleur en Bulgarie, au Danemark, en France et aux Pays-Bas, qui se situent dans le troisième quartile de la distribution.

Enfin, c'est au sein de quatorze systèmes éducatifs, dont ceux de l'Écosse, de l'Angleterre, et des provinces canadiennes de l'Ontario, de la Saskatchewan, de la Nouvelle-Écosse et de

l'Alberta, que cette composante instrumentale de comparaison entre les écoles et l'amélioration de leurs performances est la plus développée.

Ici encore, les résultats de l'ANOVA [$F(3,55) = 174,02 ; p = 0,00$] et du test de Tukey suggèrent la présence de différences statistiquement significatives entre ces quatre groupes.

Figure 6 : Composante instrumentale de comparaison externe et dispositifs internes d'amélioration des performances de l'école



4.3 Synthèse et interprétation des composantes instrumentales des systèmes d'accountability basés sur la performance

Des analyses en composantes principales, ont émergé trois composantes instrumentales des SABP qui renvoient finalement à différentes logiques de régulation : 1) une composante d'évaluation et de régulation selon la performance, 2) une composante qui se rapporte au suivi de l'évaluation des enseignants et donc, aux enjeux et incitatifs directs qui y sont attachés à ce type de dispositifs et enfin 3) une composante de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires.

La première composante renvoie ainsi à une dimension centrale des SABP, celle de l'évaluation et du suivi de la performance. Ici, la régulation s'appuie sur un ensemble d'instruments qui impliquent d'une part un suivi des résultats (des élèves et des écoles) et d'autre part, la conformité à des standards (standards de performance, standards permettant de fixer les critères d'évaluation des établissements scolaires) (Kirst 1990 cité dans McEwen, 1995). Elle repose sur

différents mécanismes clefs, dont la mise en place de dispositifs pour mesurer, consigner et suivre les résultats des élèves, et de dispositifs formels d'évaluation des établissements scolaires. La reddition de comptes échoit par ailleurs à une autorité administrative en charge du suivi des résultats des élèves et des établissements scolaires.

Cette composante instrumentale peut ainsi être interprétée à la lumière de la montée en puissance de l'évaluation des établissements scolaires et des résultats des actions des personnels éducatifs (Barrère, 2010; Felouzis & Hanhart, 2011). Elle fait écho à l'usage de l'évaluation comme instrument de pilotage et comme dispositif visant la production d'information sur la performance des élèves, mais aussi des enseignants et des établissements scolaires. Les évaluations (évaluations externes des acquis des élèves, évaluation des établissements) sont dès lors mobilisées pour orienter les comportements des acteurs locaux parallèlement à une explicitation des standards de performance à atteindre (Maroy & Voisin, 2013, p. 17-18).

Cette logique instrumentale se rapporte donc à l'une des caractéristiques centrales de la nouvelle *accountability* (Carnoy & Loeb, 2002; Fuhrman, 1999; Fuhrman & Elmore, 2004; Linn, 2000; Maroy & Voisin, 2014), centrée sur les produits et les résultats des structures et des agents éducatifs, où il s'agit de superposer ou de substituer la mesure des résultats du système au contrôle des ressources et des processus éducatifs.

L'accent est mis sur la responsabilisation des acteurs vis-à-vis des résultats des élèves sur lesquels portent les dispositifs de suivi et, dès lors, les modalités de contrôle. On peut donc supposer que la théorie de l'action sous-jacente à cette composante s'appuie sur la capacité des acteurs à ajuster ou à modifier leurs pratiques lorsqu'ils sont confrontés aux résultats de leurs actions. Les résultats des élèves permettent de renvoyer une image des actions des professionnels et des effets de leurs pratiques (Thélot, 1998 cité dans Mons, 2009; Maroy et Voisin, 2014). Thélot emploie ici la métaphore de l'« effet miroir » pour caractériser ce processus devant favoriser un agir professionnel rationnel : confrontés aux résultats de leurs élèves, les acteurs devraient être en mesure de réajuster leurs pratiques. On présuppose un engagement des acteurs envers les buts de l'organisation (performance et réussite des élèves) et leur capacité à s'engager dans un processus de réflexivité. L'effet miroir renvoie donc à « une modélisation qui s'appuie exclusivement sur une sanction symbolique » (Mons, 2009, p. 14)

Nos résultats mettent par ailleurs en exergue que cette composante instrumentale est d'ampleur différenciée au sein des systèmes éducatifs de notre échantillon. Ce sont en effet les systèmes éducatifs de certaines régions espagnoles (Madrid, Andalousie, Catalogne) et de certaines provinces canadiennes (Terre-Neuve, Alberta), ainsi que le système éducatif de l'Écosse qui se démarquent du reste de notre échantillon par l'ampleur la plus importante de cette composante en leur sein. De fait, les instruments qui la construisent y sont davantage mobilisés. À l'opposé, il apparaît que la composante d'évaluation et de régulation selon la performance est beaucoup moins développée dans les Länder allemands inclus à notre échantillon, en Grèce, en Italie, en France ou dans les communautés francophones et germanophones de Belgique.

La seconde composante extraite de nos analyses a trait à une autre dimension clef des systèmes d'*accountability* : les enjeux et incitatifs directs pour le corps enseignant. Cette composante instrumentale est en effet construite par l'ensemble des variables qui se rapportent aux conséquences matérielles et sociales articulées aux évaluations des enseignants (exception faite de la mise en place d'activités de développement professionnel). Il s'agit donc de formaliser un ensemble d'incitatifs qui font suite aux évaluations du corps enseignant, ces dernières n'étant par ailleurs pas déconnectées de l'enjeu de performance.

Pour mieux comprendre la théorie de l'action sous-jacente à cette composante qui fait des incitatifs la clef du changement, il faut se tourner vers la théorie du choix rationnel qui découle de la pensée économique néoclassique (Campbell, 2002). Elle s'appuie sur deux notions centrales pour expliquer les comportements des individus, celles d'instrumentalisme et d'analyse coût-avantage (Meadwell, 2002) (Boudon, 2004).

En effet, selon cette théorie les individus agissent suivant une logique utilitaire de calcul des coûts et des bénéfices, supposée être à la base des choix qu'ils effectuent et des comportements qui en découlent. L'action humaine est donc, selon ces cadres, « essentiellement instrumentale, et les acteurs calculent rationnellement quelles lignes d'action sont les plus susceptibles de maximiser leurs récompenses globales » (Rule, cité dans Meadwell, 2002, p. 117-118).

Dès lors, puisque les acteurs sont dotés d'une rationalité calculatrice, soumise par ailleurs aux influences des contextes institutionnels dans lesquels ils évoluent (Bates *et al.*, 1998; Frieden & Rogowski, 1996; Garrett & Lange, 1996 cités dans Campbell, 2002, p. 37), les contraintes ou les

incitatifs devraient favoriser l'orientation et l'ajustement du cours de leurs actions dans le sens désiré par l'organisation.

Les incitatifs sont au cœur de la théorie du changement organisationnel développée dans les travaux de Clark et Wilson (1961), qui s'inscrivent par ailleurs dans la lignée de ceux de Barnard (Barnard, 1938). Pour ces derniers, les systèmes d'incitatifs jouent un rôle fondamental au sein des organisations. Nécessaires à leur « bon » fonctionnement, ils permettent de garantir la coopération des acteurs et doivent dès lors être considérés comme l'une des variables centrales affectant le comportement organisationnel (Clark & Wilson, 1961) :

“The contributions of personal efforts which constitute the energies of organizations are yielded by individuals because of incentives. The egotistical motives of self-preservation and self-gratification are dominating forces; on the whole, organizations can exist only when consistent with the satisfaction of these motives, unless, alternatively, they can change these motives. The individual is always the basic strategic factor in organizations. Regardless of his history or his obligations he must be induced to cooperate, or there can be no cooperation.”
(Barnard, 1938, cité dans Clark & Wilson, 1961, p. 132)

Leurs travaux distinguent en outre différents types d'incitatifs et permettent d'éclairer la nature distincte de ceux qui construisent cette composante instrumentale des SABP. Les incitatifs matériels (Clark et Wilson, 1961) prennent la forme de récompenses (ou de sanctions) tangibles à valeur pécuniaire. Les bonus financiers et les implications en termes de changement de salaire s'inscrivent clairement dans cette catégorie.

L'attribution de responsabilités professionnelles (censées rendre la profession plus attractive) et la reconnaissance publique de la part du chef d'établissement peuvent être classées dans la catégorie d'incitatifs que Clark & Wilson (1961) nomment incitatifs téléologiques (*purposive incentives*). Ces derniers, intangibles par nature, s'appuient sur la valorisation de l'engagement et sur la participation des individus vis-à-vis des finalités et des objectifs poursuivis par l'organisation. On peut par ailleurs distinguer des incitatifs liés au statut professionnel ou social tels que les implications en termes d'avancement de carrière.

Dans cette configuration instrumentale, la reddition de comptes s'opère donc au travers des dispositifs d'évaluation des enseignants qui sont par ailleurs assortis de conséquences matérielles et sociales de natures diverses.

Nous avons vu que ce sont les incitatifs matériels qui apportent la contribution la plus importante

à la construction de cette composante. De fait, les systèmes d'éducation où ce type d'enjeux est associé aux évaluations des enseignants se sont vus attribuer les scores factoriels aux valeurs les plus élevées. Ainsi, nos résultats suggèrent que c'est majoritairement dans les pays de l'est de l'Europe tels que la Slovaquie, la Hongrie, la République tchèque, ou encore en Angleterre et en Suède que cette composante instrumentale est la plus développée.

Enfin, la dernière composante instrumentale extraite de nos analyses présente la cohérence statistique la plus faible. Elle repose sur le couplage entre mécanismes de comparaison des résultats des établissements scolaires (et indirectement, de mise en compétition de ces derniers sur cette base) et mise en place de dispositifs d'intervention en leur sein, qui impliquent la sollicitation d'experts externes d'une part, et l'opportunité pour les enseignants de participer à des activités de développement professionnel, d'autre part.

Cette composante réfère dès lors partiellement à la notion d'*accountability* de marché travaillée (Leithwood & Earl, 2000; Harris & Herrington, 2006), dans la mesure où elle repose en partie sur une logique d'action similaire qui vise à stimuler la compétition entre établissements scolaires. La compétition entre établissements est en effet induite par la publication des résultats des élèves, la comparaison des performances de ces derniers et de leurs élèves. Les suppositions défendues par les promoteurs du choix de l'école ((John E. Chubb & Moe, 1988; J.E. Chubb & Moe, 1990) pour les plus célèbres et vraisemblablement les plus cités) reposent sur l'idée que la compétition couplée à des mécanismes de reddition de comptes au public devrait *in fine* participer à l'amélioration des résultats des établissements scolaires. La publication des résultats des élèves permet donc de rendre publiques les performances des écoles et les pratiques de comparaison entre écoles de stimuler la compétition entre ces dernières.

Cette composante instrumentale se distingue toutefois de l'*accountability* de marché, car elle incorpore des mécanismes qui reposent, non pas uniquement sur l'action supposée de la concurrence pour susciter le changement, mais aussi sur des dispositifs d'intervention au sein des écoles visant à soutenir les processus (éducatifs et de gestion) dans une perspective d'amélioration de leurs résultats.

Ces dispositifs d'intervention ont été fortement recommandés par les tenants du courant de l'amélioration des écoles qui préconisent la mise en place d'un ensemble de stratégies visant à soutenir les processus éducatifs au sein des établissements scolaires. Il s'agit donc d'analyser et

d'améliorer les processus à l'œuvre, de favoriser l'adoption de pratiques éducatives et de gestion plus efficaces, d'impliquer collectivement et individuellement les acteurs dans un processus de changement et de réflexion sur leurs pratiques (Bressoux, 1995; Normand, 2006).

Au total, cette dernière composante instrumentale repose donc sur le couplage et l'alignement de divers instruments supposés vecteurs de changement (Maroy & Voisin, 2014) : d'un côté, des mécanismes de visibilisation des performances des écoles venant stimuler la compétition entre écoles et de l'autre, des stratégies d'intervention au sein des écoles qui permettent de sensibiliser les acteurs à des pratiques et à des modèles de comportements plus efficaces et qui sont alignés et couplés aux dispositifs.

D'après nos résultats, c'est dans les systèmes d'éducation de certaines provinces canadiennes telles que l'Ontario, la Saskatchewan et la Nouvelle-Écosse, et dans les systèmes d'éducation de l'Écosse, de l'Angleterre, du Pays de Galles et de l'Irlande du Nord que cette dimension est la plus développée. À l'opposé, elle est de plus faible intensité dans les régions espagnoles de la Galice, de Castille et de Léon, de la Murcia ou encore dans les Länder allemands de la Basse-Saxe (Niedersachsen) ou de l'Hesse.

En conclusion, les trois composantes extraites de nos analyses renvoient donc à des configurations instrumentales distinctes. Ces composantes des SABP se distinguent ainsi sur plusieurs aspects, dont le niveau d'action ciblé, soit l'acteur clef visé du changement, les instruments et les dispositifs centraux au cœur des logiques de régulation qui sont mises en acte par ces derniers. Si nous avons analysé ces composantes de manière indépendante et avons décrit leur ampleur dans les systèmes éducatifs de notre échantillon, ces dernières ne sont pas mutuellement exclusives, mais peuvent se combiner pour former des SABP plus complexes. Elles peuvent ainsi être articulées dans une perspective de régulation des systèmes éducatifs. Les analyses de classification hiérarchique permettent ainsi de mettre au jour comment ces composantes instrumentales des systèmes d'*accountability* basés sur la performance s'articulent au sein de types de SABP, au sein desquels sont regroupés les systèmes éducatifs de notre échantillon. Nous en présentons les résultats dans la section suivante.

SECTION 2 : types empiriques différenciés de systèmes d'*accountability* basés sur la performance

Cette section présente les résultats de l'analyse de classification hiérarchique (ACH) réalisée sur les composantes principales décrites dans les sous-sections précédentes⁵⁵. Nous décrivons premièrement la répartition de l'échantillon au sein des types empiriques qui émergent des ACH, puis les caractéristiques de ces derniers en termes essentiellement quantitatifs, c'est-à-dire sur base de leurs moyennes pour chaque composante instrumentale. Nous étayons par ailleurs la description et l'interprétation des types empiriques en tissant des liens avec la littérature. Nous proposons finalement une synthèse des résultats de ce premier volet de recherche.

4.4 Résultats des analyses de classification hiérarchique

L'analyse de classification hiérarchique nous a permis de construire une typologie exploratoire de SABP où les composantes instrumentales sont articulées et combinées selon des configurations différenciées. Cette démarche d'analyse se rapporte ainsi au deuxième postulat au cœur de ce premier volet de recherche. Nous avons en effet posé que différentes configurations d'instruments permettent de distinguer des types empiriques de SABP qui caractérisent les systèmes éducatifs de notre échantillon.

L'analyse de *cluster* aboutit à l'élaboration d'une typologie empirique qui répartit les 59 systèmes éducatifs de notre échantillon en quatre types empiriques de SABP : 1) un système d'*accountability* de faible intensité et peu instrumenté; 2) un système d'*accountability* axée principalement sur l'évaluation de la performance; 3) une *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne des performances des établissements scolaires ; et 4) une *accountability* de forte intensité prévoyant des incitatifs directs pour le corps enseignant.

Pour décrire ces types empiriques dans les sous-sections suivantes, nous procédons tout d'abord à l'examen du dendrogramme (Figure 7) qui illustre les regroupements de systèmes éducatifs selon différents types de SABP. Nous décrivons ensuite les caractéristiques des types en examinant leurs moyennes sur les composantes instrumentales. Nous étayons par ailleurs la description et l'interprétation des types empiriques de SABP en tissant des liens avec la

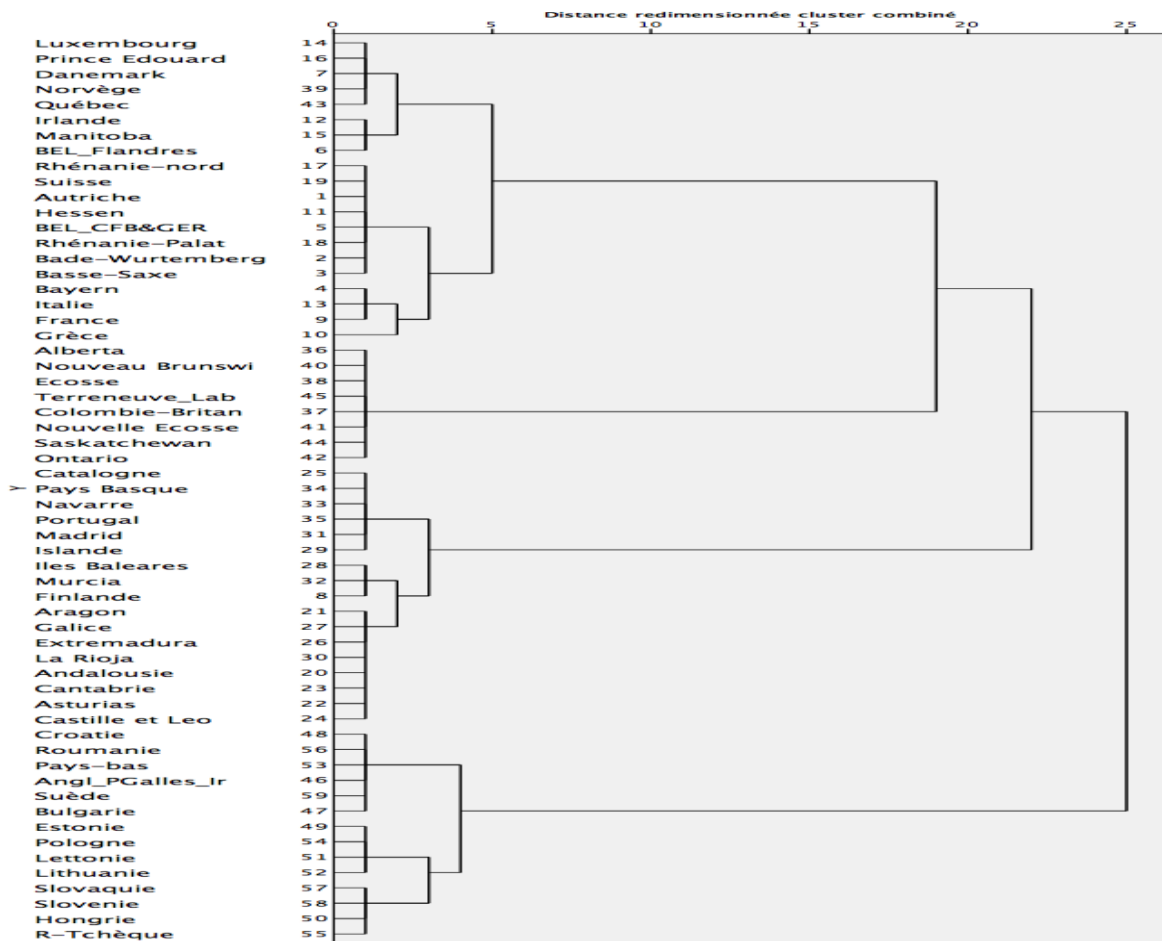
⁵⁵ Pour rappel, les scores factoriels attribués aux systèmes éducatifs pour chaque composante instrumentale sont les variables introduites pour réaliser l'analyse de classification hiérarchique.

littérature institutionnelle et empirique.

4.4.1 Répartition de l'échantillon au sein des types empiriques

Le dendrogramme produit par l'analyse de classification (Figure 7) suggère une répartition de notre échantillon en quatre groupes⁵⁶.

Figure 7 : Dendrogramme illustrant la répartition de l'échantillon au sein des types empiriques de SABP (méthode de Ward comme algorithme de regroupement et distance euclidienne au carré)



⁵⁶ Nous aurions pu opter pour une solution à sept regroupements. Nous avons fait le choix de conserver une classification à quatre groupes, car elle se prête davantage aux analyses de régressions multiples que nous réalisons dans le deuxième volet empirique de notre recherche (un découpage en sept classes rendant ce type d'analyses difficilement gérable). Toutefois, nous avons réalisé des analyses de régressions multiples en scindant le premier *cluster* en deux groupes, du fait principalement de la connaissance que nous avons des systèmes éducatifs québécois et français en ce qui a trait aux systèmes d'*accountability* qui y sont privilégiés. Les résultats des analyses de régressions multiples avec quatre groupes et cinq groupes ont été comparés. Aucune différence n'est apparue. Il ne nous semble donc pas pertinent de scinder les *clusters* en sous-groupes. Nous tâchons néanmoins, dans la description des types empiriques, de faire apparaître les nuances entre les sous-groupes de systèmes éducatifs qui apparaissent au sein des *clusters*.

Notons d'entrée de jeu que les résultats des analyses de classification opposent clairement un premier groupe de systèmes éducatifs aux trois autres groupes qui émergent des analyses. Comme nous le verrons plus avant, ce *cluster* qui compte quatorze systèmes éducatifs (soit 24 % de notre échantillon) et regroupe les pays de l'est de l'Europe, l'Angleterre, la Suède et les Pays-Bas, distingue clairement les systèmes éducatifs qui prévoient des incitatifs directs pour les enseignants du reste des systèmes éducatifs de notre échantillon. Le dendrogramme suggère en effet que les trois premiers *clusters* présentent davantage de similitudes entre eux qu'avec ce dernier.

Rappelons ici que les composantes instrumentales n'apportent pas la même contribution à la construction des *clusters* (ou types empiriques). Des analyses de variance⁵⁷ montrent en effet que c'est la composante se rapportant aux incitatifs pour les enseignants qui est la plus discriminante [F (3,55) = 66,442; $p = 0,00$], suivi par celle de l'évaluation selon la performance [F (3,55) = 35,201; $p = 0,00$], puis de compétition entre écoles et d'amélioration de leurs performances [F (3,55) = 33,670 ; $p = 0,00$]. Cette information permet de mieux comprendre les regroupements effectués par l'analyse de classification hiérarchique d'une part, les similitudes ou différences inter-*cluster*, d'autre part.

Le premier regroupement qui apparaît sur le dendrogramme, formé par les systèmes éducatifs du Luxembourg, de l'Île du Prince Édouard, de la France, de la Belgique francophone et germanophone, du Québec (entre autres) compte les effectifs les plus importants, vingt systèmes éducatifs soit 34 % de notre échantillon. Il est caractérisé principalement par une *accountability* de faible intensité, peu instrumentée. En d'autres termes les composantes instrumentales, et donc l'instrumentation sur laquelle ce type repose, sont sensiblement moins développées que dans les autres types de SABP.

⁵⁷ ANOVA à un facteur de classification où les variables indépendantes (X) sont les scores factoriels attribués aux systèmes éducatifs pour chaque composante instrumentale, le facteur de classification les types empiriques de SABP. Par ailleurs des tests de Tukey sont réalisés pour observer des différences statistiquement significatives de moyennes entre les clusters pour chaque composante instrumentale. Le caractère statistiquement significatif des différences de moyennes des types empiriques pour chaque composante est répertorié dans le Tableau 17. Puisque le postulat d'égalité des variances n'est pas respecté pour la composante de compétition entre écoles et d'amélioration de leurs performances, nous avons donc réalisé un test non paramétrique de Kruskal-Wallis. Les résultats ($\chi^2 (3) = 33.635$ $p = 0.000$) attestent de différences intergroupes pour cette composante instrumentale.

Nous observons que deux sous-groupes de systèmes éducatifs s'y distinguent néanmoins. D'un côté, les systèmes éducatifs du Luxembourg, du Danemark, de la Norvège, de l'Irlande et de la communauté flamande de Belgique sont regroupés avec les provinces canadiennes du Québec et du Manitoba. De l'autre, les *Lands* allemands, la Suisse, l'Autriche, la Belgique germanophone et francophone forment un sous-groupe distinct avec la France, la Grèce et l'Italie. Nous verrons que ces sous-groupes de systèmes éducatifs présentent en effet des distinctions en ce qui a trait à l'ampleur de certaines composantes instrumentales en leur sein.

Le deuxième regroupement est le moins peuplé. Il est constitué par huit systèmes éducatifs, soit environ 13 % de notre échantillon : le reste des provinces canadiennes (Ontario, Alberta, Nouveau-Brunswick, etc.) et l'Écosse. Il se distingue par l'articulation des composantes instrumentales d'évaluation selon la performance et de comparaison des performances des établissements scolaires qui y sont fortement développées.

Le troisième *cluster* est composé de 17 systèmes d'éducation, soit 29 % de notre échantillon. Il regroupe les régions d'Espagne avec la Finlande, l'Islande et le Portugal. Il se distingue principalement par une composante instrumentale d'évaluation selon la performance de forte ampleur et donc, par une gamme d'instruments s'y rapportant relativement étendue. Le dendrogramme suggère aussi la formation de deux sous-groupes au sein du *cluster*, avec d'un côté les régions espagnoles de la Catalogne, du Pays basque, de la Navarre, de la Murcia, des îles Baléares et de Madrid regroupées avec le Portugal et l'Islande, de l'autre, le reste des régions espagnoles et la Finlande.

Ce *cluster* se démarque quant à lui des types empiriques formés d'un côté par les systèmes éducatifs du Québec et de la France (entre autres), de l'autre par la majorité des provinces canadiennes et l'Écosse, par une composante de comparaison entre écoles moins développée. Il s'en rapproche néanmoins dans la mesure où, dans ces trois *clusters*, peu de place est accordée à la mise en œuvre d'incitatifs pour le corps enseignant, ce qui explique la proximité de ces *clusters* sur le dendrogramme.

Finalement, comme nous l'avons dit, le dernier *cluster* formé par les pays de l'est de l'Europe, l'Angleterre, la Suède et les Pays-Bas se distingue des trois groupes précédents par une composante instrumentale se rapportant aux enjeux et incitatifs directs pour les enseignants fortement développée. Aux conséquences matérielles et sociales qui font suite aux évaluations

des enseignants, il articule par ailleurs un ensemble d'instruments se rapportant d'une part à la composante d'évaluation et de régulation selon la performance, d'autre part de comparaison des résultats des établissements scolaires, ces dernières étant, comme nous le verrons, fortement développées au sein de ce type de SABP.

Avant de décrire les caractéristiques de chaque type empirique, nous proposons une première ébauche de typologie empirique. Le Tableau 16 reprend la dénomination des types empiriques de SABP que nous proposons, les regroupements de systèmes éducatifs illustrés dans le dendrogramme, les effectifs et les systèmes éducatifs qui y sont regroupés. La classification que nous proposons ne respecte pas *stricto sensu* l'ordonnement proposé par le dendrogramme. Au-delà des composantes qui permettent de discriminer les types, nous avons choisi de prendre en compte ce qui apparaît comme une gradation dans l'articulation des composantes privilégiées au sein des *clusters* pour organiser notre typologie empirique.

Tableau 16 : Première ébauche de la typologie empirique des SABP

Dénomination des types empiriques de SABP	<i>Accountability</i> de faible intensité peu instrumentée	<i>Accountability</i> axée sur l'évaluation de la performance	<i>Accountability</i> axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance	<i>Accountability</i> de forte intensité prévoyant des incitatifs directs pour les enseignants
Regroupements illustrés par le dendrogramme	1	3	2	4
Effectifs	20 (34 %)	17 (29 %)	8 (13 %)	14 (24 %)
Systèmes éducatifs	Luxembourg, Prince-Édouard, Danemark, Norvège, Québec, Irlande, Manitoba, Belgique communautaire flamande, Autriche, Bade-Wurtemberg, Basse-Saxe, Bayern, Belgique francophone, germanophone, France, Grèce, Italie, Rhénanie-Nord Rhénanie-Palatinat, Hesse, Suisse	Madrid Murcia Navarre Pays basque Portugal Islande Iles Baléares Finlande Andalousie Aragon Asturias Cantabrie Castille et Leo Catalogne Extremadura Galice La Rioja	Alberta Colombie-Britannique Écosse Nouveau Brunswick Nouvelle Écosse Ontario Saskatchewan Terre-Neuve et Labrador	Angleterre, Irlande du Nord et Pays de Galles Pays-Bas Bulgarie Croatie Roumanie Suède Estonie Hongrie Lettonie Lituanie Pologne R-Tchèque Slovaquie Slovénie

4.4.2 Caractérisation des types empiriques à partir de leurs moyennes sur les composantes

La description des caractéristiques des types empiriques réalisée dans cette sous-section est principalement organisée autour de l'examen de leurs moyennes pour chaque composante instrumentale (ANNEXE 6). Ces moyennes sont à interpréter à la lumière de la moyenne de l'échantillon total pour chaque composante⁵⁸, mais aussi à titre comparatif, c'est-à-dire en comparant les moyennes de chaque *cluster*.

La moyenne et l'écart-type des *clusters* pour chaque composante instrumentale sont présentés dans le Tableau 17. Les différences statistiquement significatives de moyennes entre *cluster* sont aussi répertoriées⁵⁹. Ces informations servent donc de base à la description des types empiriques.

⁵⁸ Moyenne de 0, écart-type de 1 pour l'ensemble de l'échantillon.

⁵⁹ Le signe < signifie que la moyenne du type répertorié en ligne est significativement moins élevée que celle des autres types dont le numéro est listé à côté du signe; le signe > signifie qu'elle est significativement plus élevée. Ainsi, lorsque nous notons pour le type 1 de SABP « < 2* » dans la colonne qui se rapporte à la composante d'évaluation selon la performance, cela signifie que la moyenne du type 1 est significativement inférieure à celle du type 2 pour cette composante instrumentale. Le caractère statistiquement significatif des différences de moyenne entre types empiriques pour chaque composante instrumentale a été testé via des analyses de variance à 1 critère de classification : les variables dépendantes sont les composantes instrumentales (X), les types de SABP le facteur de classification (voir note de bas de page précédente). Les statistiques descriptives sont placées en annexe (ANNEXE 6).

Tableau 17 : Caractéristiques des types empiriques de SABP : moyenne et écart-type pour chaque composante instrumentale et différences avec les autres types empiriques.

	Composante d'évaluation et de régulation selon la performance			Enjeux et incitatifs directs articulés à l'évaluation des enseignants			Comparaison externe et dispositifs internes d'amélioration des performances de l'école		
	Moyenne	Écart-type	Différences significatives avec types empiriques*	Moyenne	Écart-type	Différences significatives avec types empiriques*	Moyenne	Écart-type	Différences significatives avec types empiriques*
1) <i>accountability</i> de faible intensité peu instrumenté	-1,09	0,65	< 2)* < 3)* < 4)*	-0,36	0,50	> 3)* < 4)*	-0,13	0,67	> 2)* < 3)*
2) <i>accountability</i> axé sur l'évaluation de la performance	0,81	0,55	> 1)*	-0,39	0,41	> 3)* < 4)*	-0,89	0,56	< 1)* < 3)* < 4)*
3) <i>accountability</i> axée sur l'évaluation de la performance et de comparaison externe/amélioration interne des performances des établissements	0,49	0,49	> 1)*	-0,96	0,35	< 1)* < 2)* < 4)*	1,67	0,32	> 1)* > 2)* > 4)*
4) <i>accountability</i> de forte intensité prévoyant des incitatifs directs pour les enseignants	0,30	0,63	> 1)*	1,53	0,57	> 1)* > 2)* > 3)*	0,31	0,69	> 2)* < 3)*

* différence significative de moyenne au niveau 0,05.

4.4.2.1 Premier type empirique : système d'*accountability* de faible intensité et peu instrumentée

Comme nous l'avons vu, le premier type empirique compte le nombre le plus important de systèmes éducatifs. Il regroupe une partie des systèmes d'éducation d'Europe continentale (la France, la Belgique francophone et germanophone, la Suisse, etc.) avec certaines provinces canadiennes (le Québec, le Manitoba, l'Île du Prince Édouard). Il se distingue clairement par une *accountability* de faible intensité, toutes composantes confondues.

Les moyennes du *cluster* pour les trois composantes instrumentales sont en effet systématiquement inférieures à la moyenne de l'échantillon. La composante d'évaluation selon la performance y est de faible ampleur (M=-1,09; ET=0,65 [Tableau 17]), significativement inférieure à celle des autres *clusters*. Les composantes instrumentales se rapportant aux incitatifs articulés à l'évaluation des enseignants (M=-0,36; ET=0,50 [Tableau 17]) et à la compétition

entre écoles y sont aussi d'intensité faible ($M=-0,13$; $ET=0,67$ [Tableau 17]). Ce type empirique repose donc sur des configurations instrumentales peu développées.

Pour comprendre la formation des sous-groupes au sein du *cluster*, il faut par ailleurs se tourner d'une part vers la composante de comparaison entre écoles et d'amélioration de leurs performances, d'autre part vers celle d'évaluation selon la performance. Le comparatif des systèmes éducatifs par composante a en effet mis en exergue le fait que cette dernière est, dans les systèmes d'éducation du premier sous-groupe (Luxembourg, Québec, etc.), de forte ampleur. Elle est néanmoins d'intensité plus faible dans les systèmes éducatifs du second sous groupe (France, Suisse, etc.). Il en va de même pour la composante d'évaluation selon la performance, qui est d'ampleur plus forte au sein du premier sous groupe que dans le second. Ceci explique pourquoi ces groupes de pays se distinguent au sein du *cluster*.

4.4.2.2 Deuxième type empirique : *accountability* axée sur l'évaluation de la performance

Le deuxième type empirique qui émerge de nos analyses se distingue par une *accountability* axée principalement sur l'évaluation et le suivi de la performance des acteurs.

Ce *cluster* enregistre en effet la moyenne significativement la plus élevée sur cette composante ($M=0,81$; $ET=0,055$ [Tableau 17]). Les configurations instrumentales privilégiées dans les régions d'Espagne, la Finlande, le Portugal et l'Islande impliquent donc un ensemble de dispositifs tels que l'enregistrement des données de performance et de gestion par les établissements scolaires, le suivi des résultats des élèves au niveau des écoles elles-mêmes et par une autorité administrative et l'évaluation interne et externe des établissements scolaires. Il se distingue donc clairement du premier type empirique où cette logique est nettement moins prégnante.

Ce type enregistre par ailleurs des moyennes systématiquement inférieures à la moyenne de l'échantillon pour les deux autres composantes, dont la moyenne la plus faible de tous les types pour la dimension de compétition entre écoles et d'amélioration de leurs performances ($M=-0,89$; $ET= 0,56$ [Tableau 17]). Les configurations instrumentales s'y rapportant privilégient la publication des résultats des élèves. La comparaison des performances des écoles ou l'articulation d'incitatifs à l'évaluation des enseignants y sont donc peu développées.

Là encore, on peut constater la formation de deux sous-groupes au sein du *cluster*. Il faut ici se tourner vers la composante de compétition entre écoles qui explique leur formation. L'examen de la distribution des systèmes éducatifs pour cette composante suggère en effet qu'elle est de plus forte intensité en Islande, au Portugal ou encore dans la région espagnole de Navarre, situés dans le premier sous-groupe, que dans le reste des systèmes éducatifs du second sous-groupe.

4.4.2.3 Troisième type empirique : *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance

La composante d'évaluation de la performance est, comme dans le type empirique précédent, sensiblement élevée ($M=0,49$; $ET=0,49$ [Tableau 17]), la moyenne du *cluster* pour la dimension de comparaison des établissements scolaires est, en outre, la plus élevée des quatre types émergents ($M=1,67$; $ET=0,32$ [Tableau 17]). Ceci indique que la stratégie et les configurations instrumentales privilégiées dans ce type de SABP, et donc dans les systèmes éducatifs de l'Écosse, de l'Ontario, de l'Alberta, du Nouveau-Brunswick, de la Nouvelle-Écosse, de la Saskatchewan, de la Colombie Britannique et de Terre-Neuve et Labrador, reposent d'une part sur un ensemble de dispositifs visant à mesurer et à suivre les performances des acteurs, d'autre part sur des instruments permettant la mise en comparaison des résultats des établissements scolaires articulés à des dispositifs d'intervention en leur sein et supportant le développement professionnel des enseignants. Ce type empirique repose donc sur une gamme d'instruments sensiblement plus étendue que les *clusters* décrits précédemment. Au même titre que ces derniers cependant, la place accordée aux incitatifs pour les enseignants y est marginale ($M= - 0,96$; $ET=0,35$ [Tableau 17]), la moyenne sur cette composante étant en outre significativement la plus faible des *clusters*.

4.4.2.4 Quatrième type empirique : *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants

Finalement, le dernier type empirique repose sur une *accountability* de forte intensité s'appuyant sur des conséquences matérielles et sociales directement articulées aux évaluations des enseignants, toutes dimensions confondues. S'il combine composante d'évaluation selon la performance ($M=0,30$; $ET=0,63$ [Tableau 17]) et mécanismes de comparaison entre écoles couplés à des dispositifs d'intervention en leur sein ($M=0,31$; $ET=0,69$ [Tableau 17]), c'est

l'ampleur de la dimension se rapportant aux incitatifs pour les enseignants ($M=1,53$; $ET=0,57$ [Tableau 17]) qui, comme nous l'avons vu, le distingue clairement des autres types empiriques.

C'est le seul *cluster* qui enregistre systématiquement des moyennes supérieures aux moyennes de l'échantillon pour les trois composantes instrumentales, mais aussi le seul où la composante se rapportant aux incitatifs pour les enseignants est de forte intensité. La moyenne du type empirique est ici largement supérieure et significativement différente de tous les autres types de SABP.

Nous observons ici encore la formation de deux sous-groupes au sein du *cluster*. C'est ici la composante de compétition entre écoles qui explique leur formation. Elle est en effet d'une ampleur plus importante dans les systèmes éducatifs de l'Angleterre, la Suède, la Croatie, les Pays-Bas, qui forment le premier groupe, que dans les autres systèmes éducatifs du type.

4.5 Interprétation des résultats autour des types de systèmes d'*accountability* basés sur la performance

Si nous avons décrit les quatre types empiriques émergents en termes principalement quantitatifs, via leurs moyennes sur les composantes instrumentales ayant servi de base à la classification, nous étayons ici leur interprétation au moyen d'une démarche qui tisse des dialogues avec la littérature. Ceci nous permet, d'une part, de décrire les types empiriques à la lumière des configurations instrumentales qui y sont privilégiées en dialogue avec la littérature empirique et institutionnelle et d'autre part, d'appuyer les regroupements de systèmes éducatifs réalisés.

4.5.1 Type 1 : *accountability* de faible intensité et peu instrumentée

Comme nous l'avons vu, le premier type empirique se caractérise par une faible *accountability*. Les composantes instrumentales y sont de plus faible ampleur que dans tous les autres types empiriques. Les études empiriques portant sur la France, la Suisse ou la Belgique francophone, regroupées au sein de ce *cluster*, vont en effet dans ce sens. Elles mettent en exergue le fait que les systèmes d'*accountability* y sont nettement moins sophistiqués et moins instrumentés que dans d'autres systèmes éducatifs, même si la régulation par les résultats tend par ailleurs à s'y développer (Felouzis & Hanhart, 2011; Maroy, 2013; Dutercq & Cuculou, 2013, Pons, 2014 ; Maroy & Voisin, 2014; Maroy, Pons & Dupuy, 2017).

À titre d'exemple, en France, le pilotage par les résultats s'appuie majoritairement sur le

triptyque « projets [d'établissements], évaluation, contrats [d'objectifs] » (Maroy, Pons, et al., 2017; Pons, 2014). Ces instruments, mis en place progressivement dans les dernières décennies, s'inscrivent dans une volonté de pilotage par les résultats du système éducatif français et du développement d'une culture de l'évaluation en son sein (Dutercq & Cuculou, 2013). Ainsi, la Loi Organique relative aux Lois de Finances (entrée en vigueur en 2006) engage la définition d'objectifs de performance par l'État central dans le cadre de projets dévolus aux chefs d'établissement : « L'article 7 de la loi précise l'obligation de définir des objectifs, d'en associer les moyens et de poser les conditions de mesure des effets (Pouliquen, 2005 cité dans Dutercq & Cuculou, 2013). La LOLF intègre donc deux règles : une meilleure lisibilité des sommes votées par la présentation des résultats atteints, une évaluation de la politique publique » (Dutercq & Cuculou, 2013, p. 199).

Cependant, la formalisation de la régulation du système éducatif par les résultats dans ces contextes, sensiblement moins développée que dans d'autres systèmes éducatifs, n'en est qu'à ses balbutiements, malgré l'influence des conclusions de la conférence de Lisbonne ayant mis l'accent sur la nécessité pour les systèmes éducatifs européens de se doter de standards de performance (Mons, 2009).

Nous avons néanmoins constaté qu'un sous groupe, formé par le Québec, le Manitoba, les Pays-Bas, l'Irlande, se distingue par des composantes instrumentales plus développées de comparaison entre écoles et, dans une moindre mesure, d'évaluation selon la performance.. Les études sur le Québec montrent en effet qu'une *accountability* reposant sur un ensemble d'instruments centraux s'y développe, portée par la politique de « gestion axée sur les résultats » mise en œuvre dans les années 2000 (Dembélé, Goulet, Lapointe, & Deniger, 2013; Maroy, Brassard, Mathou, Vaillancourt, & Voisin, 2015, 2017; Maroy et al., 2013).

Le système d'*accountability* québécois repose en effet sur une stratégie d'alignement engageant la fixation d'objectifs de performance au niveau central et la mise en œuvre de mécanismes de contractualisation et de planification stratégique au niveau intermédiaire et local. Dans ce contexte, les écoles sont amenées à comparer les performances de leurs élèves à la lumière des standards de performance et à réaliser un suivi pour les améliorer, les commissions scolaires (CS - instance de régulation intermédiaire) opèrent un suivi serré des performances des écoles dont elles ont la charge. Les résultats des écoles agrégés au niveau CS constituent en effet la base à

partir de laquelle la réalisation des objectifs de performance des CS (fixés par le ministère et soumis à une dimension contractuelle) est évaluée par le ministère de l'Éducation. Par ailleurs, la reddition de comptes à l'autorité hiérarchique (des établissements vers les CS, des CS vers le ministère) a lieu par ailleurs sur une base régulière. Un ensemble de dispositifs de gestion des données de performance est également mis à leur disposition afin qu'elle puisse opérer un suivi des résultats des élèves, favorisant la comparaison des performances des écoles aux performances régionales et nationales et entre écoles d'une même commission scolaire (Maroy et al., 2015; Maroy et al., 2013).

Si nos résultats tendent à mettre en exergue cette distinction vers des configurations instrumentales plus développées dans le premier sous-groupe dont fait partie le Québec, et vont donc en partie dans le sens des études empiriques sur ce système d'éducation, un regroupement à quatre classes ne prend pas en compte ces divergences, mais plutôt les similitudes entre les systèmes éducatifs du *cluster*⁶⁰. Nous retenons donc que ce type empirique se caractérise principalement par des configurations instrumentales moins développées que dans les autres types de SABP, observation qui n'est en outre pas invalidée par les nuances apportées par la formation des sous-groupes.

4.5.2 Type 2 : *accountability* axée sur l'évaluation de la performance

Le deuxième type empirique se caractérise quant à lui par une *accountability* axée principalement sur l'évaluation et le suivi de la performance, tel qu'en atteste l'ampleur de cette composante. Nos résultats suggèrent en effet que dans ce type de SABP, le pilotage du système repose sur un ensemble d'instruments visant à consigner et à suivre les résultats des élèves et des établissements scolaires, au niveau des écoles d'une part, mais aussi au niveau des autorités hiérarchiques.

En effet, dans les années 2000, l'Espagne a été témoin de nombreuses réformes axées sur la mise en place de dispositifs d'évaluation des établissements et des acquis des élèves, ayant pour visées l'amélioration des performances des élèves et, plus largement, la qualité de l'éducation (Verger & Curran, 2014). Dans ce sens, l'évaluation externe des établissements est devenue obligatoire dans l'ensemble des régions autonomes d'Espagne à la suite de la Loi nationale de 2006 (Glenn

⁶⁰ Nous reviendrons sur cette question dans la discussion des résultats de notre recherche doctorale.

& Galan, 2011, p. 182) bien que les procédures d'évaluation des établissements varient en fonction des régions (Eurydice, 2015). La Loi donne par ailleurs aux régions espagnoles le pouvoir d'octroyer davantage d'autonomie aux écoles dans la gestion du personnel éducatif. Ces dernières peuvent dès lors définir des standards spécifiques de qualification pour les enseignants qu'elles soumettent aux régions en charge de leur embauche. Les établissements scolaires doivent par ailleurs mettre en place des plans de gestion (*management plans*) permettant de fixer leurs priorités qu'elles doivent rendre publics. Des évaluations sont par ailleurs organisées sur la base d'un échantillon d'écoles afin d'opérer un contrôle sur le niveau d'acquisition des élèves à l'entrée du secondaire inférieur de la scolarité obligatoire (Fuentes, 2009).

En Finlande, l'évaluation externe des établissements est organisée par les autorités locales, en l'occurrence les communes, qui décident, le cas échéant, de mettre en place ce type de dispositif dans les écoles dont elles sont responsables (Eurydice, 2015; Halinen, 2006; Halinen, Kauppinen, & Yrjölä, 2006). Le système finlandais n'a par ailleurs pas mis en place d'exams nationaux avant la fin de la scolarité obligatoire. Les conceptions de l'évaluation des élèves au sein du système éducatif finlandais reposent avant tout sur une évaluation formative, basée sur le développement des compétences des élèves. Cela n'implique pas l'absence du suivi de la progression des apprentissages des élèves, ni la fixation de standards au niveau national. En effet, des critères d'évaluation pour les disciplines de l'enseignement fondamental sont inclus dans le plan national finlandais de 2004. Ils « donnent des directives pour les écoles et les enseignants, en définissant les niveaux de compétences et de connaissances que les élèves doivent maîtriser à l'issue de leur neuvième année d'école de base » (Halinen, Kauppinen & Yrjölä, 2006, p. 81). Par ailleurs les évaluations nationales des résultats de l'apprentissage sont organisées selon des procédures d'échantillonnage (Halinen, 2006, p. 77). Les établissements sélectionnés ont l'obligation légale d'y participer. Un retour sur l'évaluation est transmis aux écoles deux mois après leur participation et permet « de situer la manière dont les établissements se sont acquittés des objectifs nationaux et les écarts par rapport à la moyenne nationale » (Halinen, Kauppinen & Yrjölä, 2006, p. 111). Le ministère de l'éducation finlandais ayant pour fonction « d'évaluer le succès de la politique éducative et de chercher à atteindre l'égalité et la qualité dans l'éducation » (Halinen, 2006, p. 77).

Au total, il apparaît que ce type empirique se caractérise par la grande importance accordée à l'amélioration des apprentissages des élèves, des résultats des établissements scolaires et du

système au sens large. Il repose sur un ensemble d'instruments mobilisés dans ce sens. Comme nous l'avons vu, la compétition entre écoles de même que l'articulation d'incitatifs aux évaluations des enseignants y sont peu valorisées. En Finlande, la non publication des résultats des écoles a fait l'objet d'un consensus national. La production d'un palmarès des écoles est proscrite et la comparaison des écoles sur base de leur performance est fortement décriée (Eurydice, 2009).

Cependant, nos résultats ont également mis en exergue le fait qu'un sous-groupe de systèmes éducatifs formé entre autres par l'Islande et le Portugal, se distingue sur ce point. Au sein de ce dernier, les configurations instrumentales visant à favoriser la compétition entre établissements scolaires au moyen de stratégies de visibilité des résultats des élèves y sont donc davantage développées qu'ailleurs.

4.5.3 Type 3 : *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance

Le troisième type empirique, formé par la majorité des provinces canadiennes et l'Écosse, repose quant à lui sur des configurations instrumentales orientées d'une part, par une logique de comparaison et d'amélioration des performances des écoles et, d'autre part, sur l'évaluation et le suivi de la performance. Il repose donc sur une gamme d'instruments sensiblement plus étendue que les types empiriques précédents où la place accordée au suivi des résultats des acteurs (élèves, enseignants, établissements scolaires) et à leur amélioration est centrale, de même que la comparaison des résultats des écoles (vis à vis des résultats des autres établissements scolaires, à la lumière des performances nationales et régionales). Il s'agit donc d'effectuer un suivi serré des résultats des élèves, mais aussi de miser d'une part, sur des dispositifs externes de mise en comparaison de ces derniers et, d'autre part, sur des dispositifs internes aux établissements visant l'amélioration de leurs performances.

Les systèmes éducatifs de l'Ontario et de l'Écosse, qui reposent sur des systèmes d'*accountability* qui présentent de fortes similitudes, sont emblématiques de l'importance accordée au suivi des résultats des élèves et à l'amélioration des performances des établissements scolaires. En Ontario, à partir des années 1990, des réformes introduisant un ensemble d'instruments dans une visée de contrôle et d'amélioration des performances du système éducatif

sont mises en œuvre (Chan et al., 2007). Les gouvernements successifs créent diverses instances (Office de la qualité et de la responsabilité en éducation - OQRE, Ordre des enseignants et enseignants de l'Ontario) et dispositifs visant d'une part, à centraliser le système éducatif ontarien, d'autre part, à améliorer son efficacité. Ces dispositifs reposent sur le suivi des résultats des élèves au niveau des établissements et par les conseils scolaires (instances de régulation intermédiaire), sur la fixation de standards de performances et l'obligation, pour les établissements scolaires, de mettre en œuvre des plans d'amélioration devant soutenir la construction des capacités (*capacity building*) (Jaafar & Anderson, 2007).

L'OQRE est en outre chargée de récolter, d'analyser et de diffuser de l'information sur la qualité du système éducatif ontarien. Elle orchestre les évaluations à grandes échelles des acquis des élèves et la publication des résultats des élèves à l'échelle provinciale. Le suivi des performances des établissements scolaires en référence des *benchmarks* nationaux ou régionaux, mais aussi en référence aux performances des autres écoles est ainsi fortement encouragé. L'OQRE exigeant en outre que les écoles prennent en compte les résultats provinciaux et les données locales relatives aux besoins des élèves et des écoles dans leur plan d'amélioration (Chan, Fisher & Rubenson, 2007).

Par ailleurs, un ensemble de consultants et de « coachs » travaillent activement avec les conseils scolaires, les établissements et les chefs d'établissements dans une perspective d'amélioration des écoles. Le développement professionnel, ou « perfectionnement professionnel » des enseignants (MEO, 2010) est partie intégrante du système d'évaluation du rendement du personnel enseignant. Il prévoit en outre l'élaboration d'un « Plan annuel de perfectionnement des enseignants » révisé et mis à jour chaque année avec les directions d'établissements. Le développement professionnel des enseignants repose aussi sur un ensemble d'instances comme les « communautés d'apprentissage professionnel », supposées favoriser « un rendement élevé chez le personnel enseignant » en stimulant la collaboration et mettant « l'accent sur les résultats » (MEO, 2010, p. 7). Les dispositifs institués combinent donc « à la fois une pression en faveur d'une obligation de résultats et diverses pratiques de soutien des districts et des établissements afin qu'ils relèvent les défis associés à cette obligation » (Lessard & Desjardins, 2011, p. 89).

En Écosse, dès les années 1980-1990, la politique d'assurance qualité et d'évaluation (*Quality Assurance and Evaluation*) introduit l'obligation pour les écoles écossaises de mettre en place des dispositifs d'évaluation interne, encadrée par un ensemble de procédures contraignantes où standards qualitatifs et quantitatifs régissent l'évaluation des processus éducatifs et de gestion des écoles (Ozga & Grek, 2012). Les plans de développement des écoles (*Development planning*) comprenant des outils statistiques et permettant, entre autres, d'évaluer les performances des écoles sont aussi progressivement institués (Croxford, Grek, & Shaik, 2009). Cette approche, basée sur l'auto-évaluation et l'amélioration des performances des établissements scolaires, constitue ainsi la pierre angulaire du système d'*accountability* écossais. L'inspection écossais joue ici un rôle d'expertise clef dans le processus de suivi et d'amélioration des performances des écoles. Il est chargé du contrôle du rendement scolaire et des statistiques produites par les écoles, de la mise en place de mesures de soutien au processus d'auto-évaluation et d'amélioration des performances des établissements, de même que du support et de l'intervention auprès des gestionnaires scolaires et des équipes éducatives. En outre, la compétition entre écoles est encouragée au sein du système éducatif écossais par l'*Education (Scotland) Act* dès 1981. La législation octroie le droit aux parents de demander l'accès aux écoles à l'extérieur de la zone géographique dont ils dépendent habituellement. L'*Education school Act* de 1992 introduit par ailleurs la publication des résultats des élèves permettant l'accès à l'information sur les performances des écoles.

4.5.4 Type 4: *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants

Finalement, le dernier type empirique qui émerge des ACH est celui où l'*accountability* est la plus développée. S'il combine les trois composantes instrumentales, c'est en outre l'articulation d'un ensemble d'incitatifs aux évaluations des enseignants qui le distingue des autres types de SABP.

Nombreux sont les travaux qui se sont intéressés à l'*accountability* mise en œuvre dans le système éducatif anglais, regroupé sous ce type empirique, qualifié de système d'*accountability* dure du fait d'une large gamme d'instruments et de conséquences matérielles et sociales importantes pour les acteurs. L'*accountability* éducative anglaise repose sur la centralisation du contrôle, la mise en place par l'État central d'un ensemble d'indicateurs de performance, ainsi

que sur le déploiement d'un appareil sophistiqué de production et de recueil de données sur le fonctionnement du système éducatif (Ozga, 2009). Elle prévoit en outre un suivi continu des résultats des écoles (Ozga, 2009; Ozga & Grek, 2010). Il est attendu des différents niveaux d'action (central, local), dont plus particulièrement les établissements scolaires, qu'ils développent leur capacité à gérer et à utiliser les données en vue de l'amélioration de leurs résultats. Un processus d'évaluation externe des établissements y est par ailleurs orchestré par le bureau pour les standards en éducation (*Office of Standards in Education, Children's Service and Skills* - OFSTED) dès les années 1990 (Dumay, 2010). Reposant sur un dispositif contraignant qui implique l'intervention des Services de l'OFSTED au sein des établissements scolaires, il vise l'évaluation des résultats des élèves, des processus éducatifs, des pratiques de gestion et du *leadership* des cadres éducatifs (Rosenthal, 2004). Au milieu des années 2000, il est couplé à des procédures d'auto-évaluation empruntées au modèle écossais (Dumay, 2010). La compétition entre écoles y est par ailleurs très présente dans un contexte de marché scolaire fort développé (Ball, 2003), nourri, entre autres, par la publication des résultats des établissements scolaires et la comparaison des performances des établissements scolaires.

Si la littérature sur les systèmes éducatifs des pays de l'Est de l'Europe aussi regroupés sous ce type de SABP est moins accessible, les travaux de Berényi, Bajomi et Neumann (Berényi, Bajomi, & Neumann, 2013) sur l'*accountability* au sein du système éducatif hongrois montrent cependant l'importance accordée à la production d'information dans la rhétorique politique postcommuniste hongroise, ainsi qu'à « l'évaluation, la mesure des performances et le développement de la qualité » (Berényi, Bajomi & Neumann, 2013, p. 76). Les données de l'enquête INES (*International Indicators of Education Systems* – OCDE) confirment par ailleurs la mise en œuvre d'incitatifs financiers entre autres pour les enseignants tchèques, hongrois, polonais, slovaques et suédois (Wößmann, 2011, p. 407). Le rapport d'Eurydice sur l'évaluation des enseignants (2008) met quant à lui de l'avant l'implémentation de modalités de contrôle de la performance des enseignants « associées à des incitations financières et à l'organisation de carrières différenciées » (Eurydice, 2008, p. 63) en Angleterre, en Bulgarie ou encore en Suède. Cette littérature va dès lors dans le sens de nos résultats qui suggèrent que ce type de SABP se distingue plus particulièrement par l'importance qui y est accordée aux incitatifs pour le corps enseignant.

En guise de synthèse, si la littérature nous a permis d'appuyer les regroupements de systèmes éducatifs qui émergent de nos analyses, nous pouvons mettre en avant le fait que les types de SABP reposent au total sur des configurations instrumentales différenciées des logiques de régulation distinctes qui sont mises en acte par des dispositifs divers :

- Le type 1 se caractérise ainsi par une *accountability* de faible intensité et peu instrumentée. Les composantes instrumentales y sont de faible ampleur, la gamme d'instruments mobilisée est sensiblement moins étendue que pour les autres types de SABP, même si cela n'implique pas l'absence de la régulation par les résultats dans les systèmes éducatifs qui y sont regroupés. La littérature sur les systèmes éducatifs regroupés sous ce type suggère que la régulation par les résultats y est tendanciellement peu formalisée. Des modes de régulation y sont bien à l'œuvre (dispositifs d'évaluation et de suivi des résultats des élèves et des établissements principalement), mais ils demeurent en général peu articulés les uns aux autres. On peut donc en déduire que la régulation par les résultats reste de faible intensité dans ce type de SABP.
- Le type 2, qui se définit par une *accountability* axée sur l'évaluation de la performance, repose sur la mise en place d'un ensemble d'instruments mobilisés à cette fin. La régulation du système s'exerce ainsi principalement au moyen de dispositifs d'évaluation et de mécanismes de suivi des résultats des élèves et des établissements scolaires. La reddition de compte opère principalement des établissements scolaires à l'autorité administrative en charge du suivi des résultats des écoles et de leurs élèves. On peut donc faire l'hypothèse que l'*accountability* repose sur la conjugaison de formes organisationnelles hiérarchiques et managériales du fait de l'importante marge de manœuvre qui est laissée aux établissements scolaires placés cependant sous le contrôle des autorités administratives.
- Le type 3 conjugue composantes instrumentales d'évaluation selon la performance et de comparaison externe couplée à des dispositifs internes d'amélioration des performances des établissements scolaires. Il repose sur un ensemble de mécanismes de comparaison des performances entre écoles, combinés à des dispositifs d'évaluation et de contrôle des résultats des élèves, et de reddition de comptes sur cette base à l'attention des autorités administratives et du public, dispositifs visant à favoriser et à accompagner les processus

de changement au sein des établissements scolaires (Maroy & Voisin, 2014). Il articule donc pression à la performance et soutien aux acteurs locaux dans une perspective d'amélioration des résultats des élèves. Il repose par ailleurs sur une gamme d'instruments sensiblement plus étendue que les premiers types de SABP décrits.

- Enfin, le type 4 repose quant à lui sur une *accountability* de forte intensité articulant plusieurs configurations instrumentales. Il combine dès lors un ensemble de dispositifs de suivi des performances des élèves et des établissements scolaires à une multiple reddition de comptes qui s'exerce d'une part aux autorités administratives et, d'autre part, à destination du public. Si l'accent y est mis à la fois sur la pression à la performance et sur les mécanismes de compétition entre écoles, c'est plus particulièrement les conséquences sociales et matérielles attachées aux évaluations des enseignants qui le distinguent des autres types empiriques de SABP.

4.6 Conclusion

Ce premier volet empirique de recherche nous a ainsi permis de caractériser les SABP dans les systèmes d'éducation d'Europe et du Canada à partir des instruments qui y sont privilégiés. Nos premières séries d'analyse (ACP), nous ont permis de faire émerger trois composantes instrumentales des SABP, construites par des instruments distincts visant toutefois un objectif commun, celui de l'amélioration des résultats des élèves.

Notre seconde série d'analyse a permis d'observer dans quelle mesure ces composantes se combinent au sein de types empiriques de SABP. Ces types se distinguent ainsi par des gammes d'instruments plus ou moins développées, des modalités de suivi et de contrôle des performances des élèves, ainsi que sur des mécanismes de redditions de comptes distincts. Les composantes des SABP, mises en exergue par notre première série d'analyses, y sont donc plus ou moins prégnantes. Selon les cas, elles sont de faible ampleur ou articulées les unes aux autres pour mettre en acte des logiques de régulation qui s'appuient sur un ensemble d'instruments clefs censés favoriser le changement et permettre l'amélioration des performances du système.

Si ce premier volet de recherche nous a permis de différencier les SABP dans les systèmes éducatifs de notre échantillon, la question centrale au cœur de notre thèse a trait aux effets des SABP sur les degrés d'efficacité et d'équité scolaires. Les SABP permettent-ils de rendre les

systèmes éducatifs plus efficaces et plus équitables ? Observe-t-on des relations différenciées entre SABP, efficacité et équité scolaire en fonction d'une part des composantes instrumentales, d'autre part des types distincts de SABP. Notre deuxième volet de recherche permet d'apporter des éléments de réponse à ces questions.

CHAPITRE 5 : QUELS SONT LES EFFETS DES SYSTÈMES D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE SUR L'EFFICACITÉ ET L'ÉQUITÉ SCOLAIRES ?

Le premier volet empirique de recherche, dont nous venons de présenter les résultats, poursuivait le double objectif analytique et descriptif de caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada. Sur cette base, notre deuxième volet empirique vise à répondre à la question centrale des effets des SABP sur les degrés d'efficacité et d'équité scolaires.

Puisque l'hypothèse centrale au cœur de ce volet de recherche est que les SABP influencent de manière différenciée les degrés d'efficacité et d'équité scolaires en fonction des composantes instrumentales privilégiées et des types qu'ils revêtent, nous avons réalisé une série d'analyses permettant de tester cette hypothèse et de répondre à la question de recherche suivante : les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables ?

La première série d'analyses, dont nous présentons les résultats dans une première section, vise à interroger les effets des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon. Nous mobilisons ici la première série de variables originales créée dans notre premier volet de recherche se rapportant aux trois composantes instrumentales des SABP décrite en amont : 1) une composante instrumentale d'évaluation et de régulation selon la performance ; 2) une composante reposant sur la mise en place d'incitatifs directs pour les enseignants ; 3) une composante instrumentale de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires. Ces variables prennent la forme de scores factoriels (variables continues - échelle d'intervalle), se rapportant à chaque composante des SABP et permettant d'en mesurer l'ampleur dans les systèmes éducatifs de notre échantillon. Il s'agit donc tout d'abord d'observer dans quelle mesure ces composantes instrumentales, analysées indépendamment les unes des autres, altèrent, le cas échéant, les degrés d'efficacité et d'équité scolaires des systèmes d'éducation d'Europe et du Canada.

La seconde série d'analyses, dont les résultats sont discutés dans une deuxième section, vise à analyser les effets des types empiriques de SABP sur l'efficacité et l'équité scolaires des systèmes éducatifs de notre échantillon. Nous avons en effet dégagé des analyses de classification hiérarchique quatre types de SABP : 1) une *accountability* de faible intensité peu

instrumentée ; 2) une *accountability* axée sur l'évaluation de la performance ; 3) une *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne des performances des établissements scolaires ; 4) une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants. Nous avons donc créé une variable nominale à quatre modalités référant à chacun des types empiriques construits, au sein desquels les systèmes éducatifs d'Europe et du Canada sont répartis. Il s'agit donc d'analyser dans quelle mesure les types de SABP altèrent, le cas échéant, potentiellement de manière différenciée, les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon.

À toutes fins utiles, le Tableau 18 ci-dessous reprend les principaux éléments de ce volet de recherche.

Tableau 18 : Rappel, tableau synoptique du volet empirique 2 de recherche

	Volet empirique 2
Questionnement central de la thèse ?	Quels sont les effets des SABP sur le degré d'efficacité et d'équité scolaires des systèmes d'éducation?
Sous-question de recherche	Les SABP permettent-ils de rendre les systèmes éducatifs plus efficaces et plus équitables?
Objectifs	<u>Analytique</u> : 1) analyser les effets des SABP sur les degrés d'efficacité et d'équité scolaires
Hypothèse	2) Effets différenciés des SABP (composantes, types empiriques) sur l'efficacité et l'équité scolaires (égalité de résultats, égalité des chances de résultats et égalité des acquis de base)
Concepts clefs	SABP, efficacité scolaire, équité scolaire
Principales méthodes d'analyse des données	Analyses bivariées (corrélations, régression simple, analyse de variance) Analyses multivariées (régressions multiples)

Ce chapitre est donc divisé en deux sections. Dans la première, nous présentons les résultats de la première série d'analyses visant à étudier les effets des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada. Dans la seconde partie, nous présentons les résultats des analyses qui permettent de mettre en exergue les effets des types empiriques de SABP sur ces mêmes dimensions.

SECTION 1 : Composantes instrumentales des systèmes d'*accountability* basés sur la performance, efficacité et équité scolaires des systèmes éducatifs d'Europe et du Canada : quelles relations, quels effets ?

Cette partie vise donc à mettre en exergue et à discuter les effets des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires⁶¹ des systèmes éducatifs de notre échantillon. Elle est composée de quatre sous-sections. Dans la première, nous réalisons tout d'abord une brève synthèse descriptive des degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada. Nous montrons la variation importante qui existe entre ces derniers en ce qui a trait aux degrés d'efficacité scolaire et d'équité scolaire (inégalités de résultats, des chances de résultats et des acquis de base. Nous présentons par ailleurs les résultats des analyses de corrélations entre les composantes instrumentales des SABP et les indicateurs d'efficacité et d'équité scolaires, mais aussi les variables de contrôle intégrées à nos analyses (dépenses publiques en éducation, variation du niveau socioéconomique des élèves, stratification des systèmes scolaires). Nous discutons par la suite des effets bruts des composantes instrumentales des SABP sur l'efficacité et l'équité scolaires des systèmes éducatifs de notre échantillon. Enfin, nous présentons les résultats des modèles de régressions multiples permettant de saisir les effets nets des composantes instrumentales des SABP sur ces dimensions, c'est-à-dire leurs effets une fois que des facteurs clefs (les variables de contrôle) sont intégrés aux analyses et nous permettent d'adopter un raisonnement « toutes choses étant égales par ailleurs ».

5.1 Analyse descriptive comparative des degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada

En préambule de nos analyses, il nous semble essentiel de dresser un portrait des systèmes éducatifs de notre échantillon en termes de degrés d'efficacité et d'équité scolaires. L'analyse descriptive commentée brièvement dans cette sous-section met en effet en exergue de fortes variations entre les systèmes éducatifs d'Europe et du Canada pour l'ensemble des indicateurs mobilisés⁶². Ce descriptif est centré sur la discipline de référence du PISA 2012, les

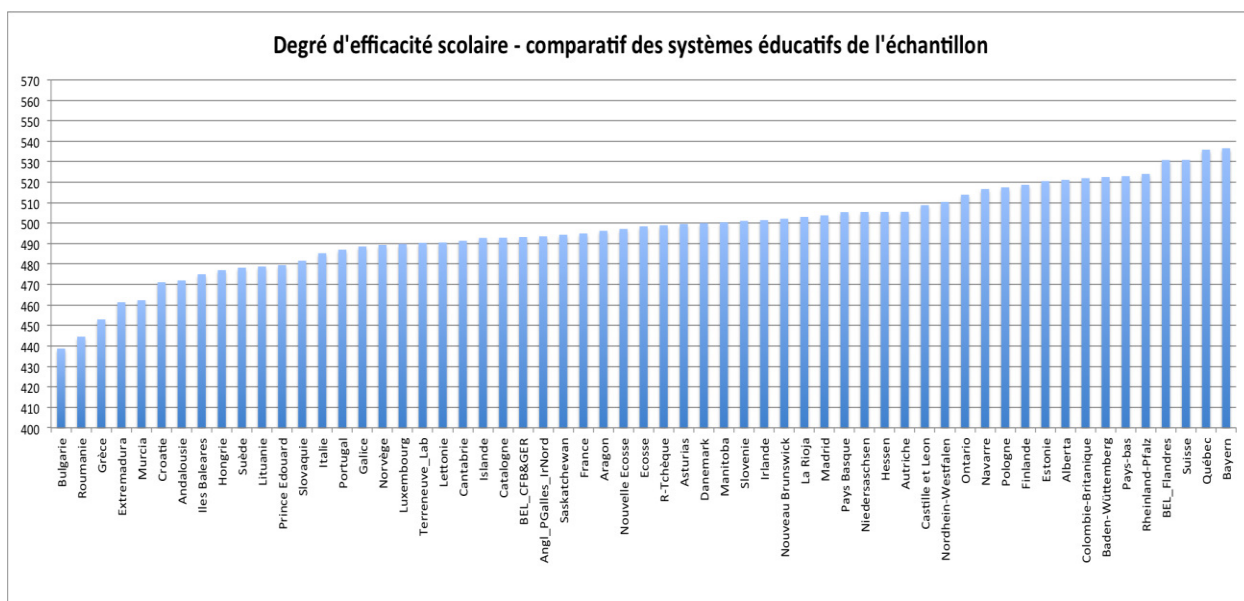
⁶¹ Voir le CHAPITRE 3 : METHODOLOGIE pour la construction des indicateurs d'efficacité et d'équité scolaire.

⁶² Pour rappel nous mobilisons un indicateur d'efficacité scolaire, 1) le niveau de performance moyen des élèves dans chaque systèmes éducatif comme indicateur d'efficacité scolaire et trois indicateurs d'équité : 2)

mathématiques, par ailleurs mobilisées dans nos analyses subséquentes⁶³. Toutefois, à toutes fins utiles, les informations pour les indicateurs en lecture et en sciences sont aussi reportées dans le texte.

Si l'on s'intéresse tout d'abord aux degrés d'efficacité scolaire dans la discipline de référence (Figure 8), on observe de fortes disparités entre les systèmes éducatifs de notre échantillon (Maths : $M=497$; $ET=21,22$; Lecture : $M=535$; $ET=20,29$; Sciences : $M=513$; $ET=23,49$). Ainsi, l'écart de performance en mathématiques est d'environ 100 points de moyenne entre les systèmes éducatifs les plus performants tel que le *Land* allemand du Bayern, et les moins performants comme la Bulgarie.

Figure 8 : Degré d'efficacité scolaire (mathématiques) – comparatif des systèmes éducatifs de l'échantillon

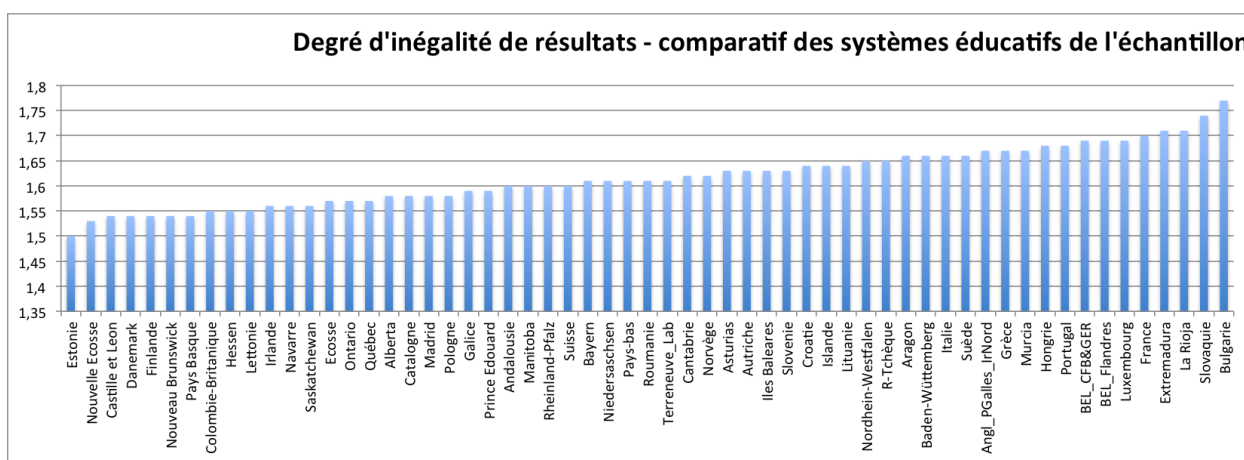


l'indicateur d'inégalité de résultats renvoie au rapport de performance entre les élèves situés dans le dernier décile de la distribution des scores et dans la premier décile, soit entre les élèves les plus et les moins performants au sein de chaque système éducatif ; 3) un indicateur d'égalité des chances de résultats qui prend la forme du coefficient de détermination (R^2) d'un modèle de régression linéaire où la variable dépendante (Y) est la moyenne des scores dans chaque matière, la variable indépendante (X) la moyenne de l'indicateur de milieu socioéconomique (IESCS dans la base de données du PISA) ; 4) un indicateur d'égalité des acquis de base qui a trait à la proportion d'élèves sous le niveau de compétences 2 du PISA dans chaque système d'éducation.

⁶³ Du fait des problèmes d'asymétrie des distributions pour les indicateurs d'efficacité scolaire en sciences, d'inégalité de résultats en science et en lecture que la transformation logarithmique appliquée aux données n'a pas permis de palier, les résultats des analyses ne sont présentés que pour les mathématiques.

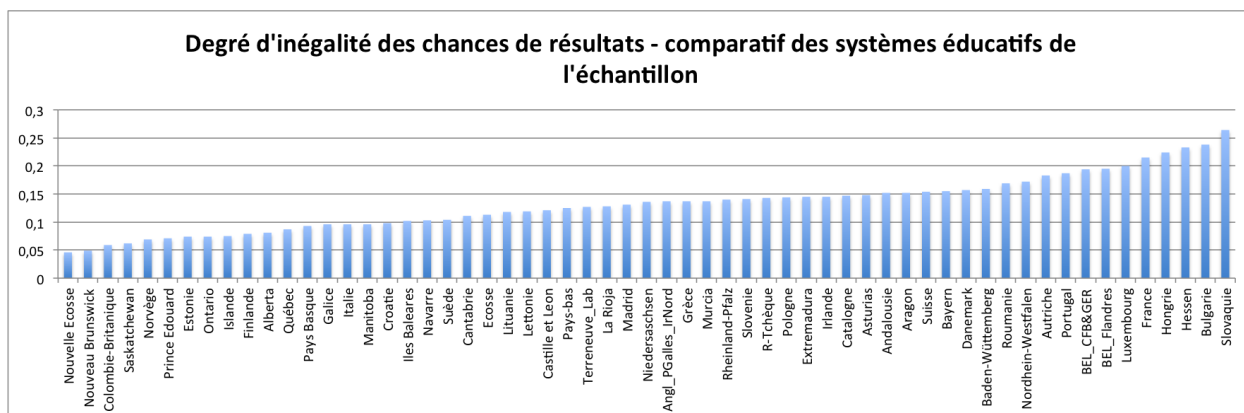
Si l'on s'intéresse maintenant à l'équité scolaire dans cette même discipline (Figure 9), on observe aussi premièrement des différences importantes dans le degré d'inégalité de résultats des systèmes éducatifs d'Europe et du Canada (Maths : $M=1,62$; $ET=0,06$; Lecture : $M=1,64$; $ET=0,10$; Sciences: $M=1,60$; $ET=0,07$;). En mathématiques, il se caractérise par un rapport de performance entre les élèves les plus et les moins performants d'environ 1,5 dans les systèmes éducatifs où ce dernier est le plus faible comme l'Estonie, de 1,77 en Bulgarie où il est le plus fort.

Figure 9 : Degré d'inégalité de résultats (mathématiques) – comparatif des systèmes éducatifs de l'échantillon



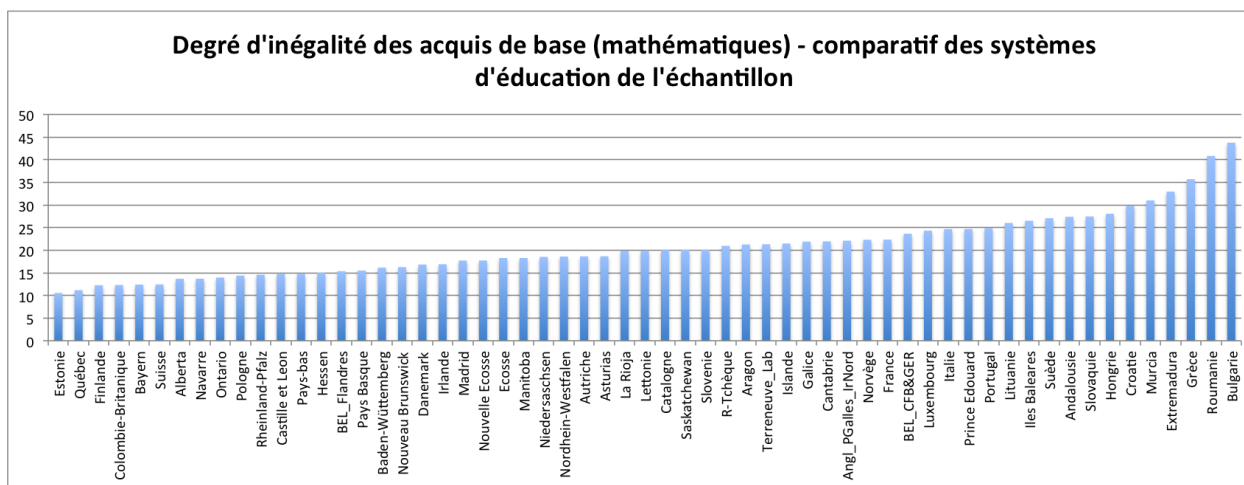
La variation inter-systèmes éducatifs en termes de degré d'inégalité des chances de résultats (Figure 10) est aussi remarquable (Maths : $M=0,14$; $ET=0,04$; Lecture : $M=0,12$; $ET=0,04$; Sciences : $M=0,13$; $ET=0,05$). Dans les systèmes d'éducation où la détermination socioéconomique des performances est la plus faible, tels que ceux des provinces canadiennes de la Nouvelle-Écosse ou du Nouveau-Brunswick, la part de variance des scores en mathématiques expliquée par le milieu socioéconomique des élèves est d'environ 5 %; elle avoisine les 25 % en Slovaquie où le degré d'inégalité des chances de résultats en mathématiques est le plus fort.

Figure 10 : Degré d'inégalité des chances de résultats (mathématiques) – comparatif des systèmes éducatifs de l'échantillon



Finally, the finding is the same if one is interested in the degree of inequality in basic skills (Maths : $M=20,75$; $ET=6,95$; Lecture : $M=17,13$; $ET=6,19$; Sciences : $M=15,33$; $ET=6,30$) (Figure 11). The proportions of students who have not reached, at 15 years, the level of competence 2 at PISA in mathematics are in fact below 15 % in the educational systems where the degree of inequality in basic skills is the lowest, such as in Estonia, whereas they exceed 40 %, in those where it is the highest, such as in Bulgaria.

Figure 11 : Degré d'inégalité des acquis de base (mathématiques) – comparatif des systèmes éducatifs de l'échantillon



Afin de compléter la description du profil des systèmes éducatifs de notre échantillon en termes d'efficacité et d'équité scolaires, nous en croisons deux aspects clefs : le degré d'efficacité scolaire et le degré d'inégalité de résultats (Figure 12⁶⁴). Ceci nous permet d'observer que certains systèmes éducatifs sont à la fois très performants et présentent par ailleurs un degré de détermination socioéconomique des résultats des élèves très faible. À l'opposé, d'autres sont peu performants et présentent un niveau de détermination socioéconomique des performances des élèves très fort. Deux systèmes qui s'opposent nettement sur ces deux axes, la Colombie-Britannique (identifiée en rouge sur la Figure 12) et la Bulgarie (identifiée en noir sur la Figure 12) permettent d'illustrer ces différences assez radicales de profil sur ces deux axes.

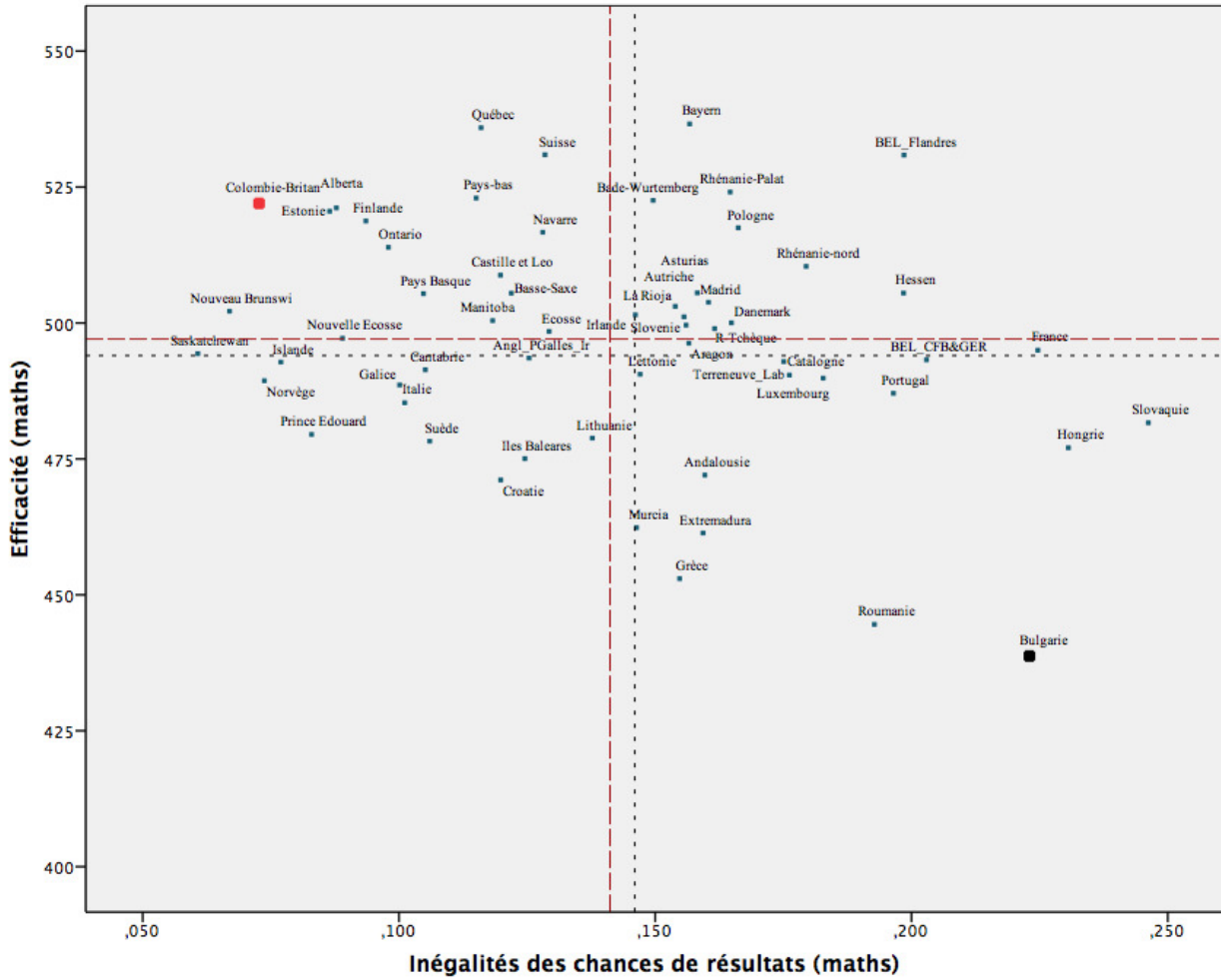
La Colombie-Britannique est en effet caractérisée à la fois par des performances moyennes parmi les plus élevées de notre échantillon ($M=522$) et un degré d'inégalité des chances de résultats, en d'autres termes, par une proportion de la variation des scores des élèves expliquée par leur niveau socioéconomique parmi les plus faibles ($M=0,07$, soit 7 %). La Bulgarie, à l'opposé, se distingue par des performances moyennes parmi les plus faibles ($M=439$) et un degré d'inégalités des chances de résultats parmi les plus élevés de notre échantillon ($M=0,22$, soit 22 % de la part des scores en lecture expliquée par le niveau socioéconomique des élèves).

Si nous prenons en compte l'ensemble des indicateurs mobilisés, la Colombie-Britannique se démarque par son système éducatif à la fois performant et très équitable : des performances moyennes élevées, de faibles différences de performances entre élèves, couplées à une détermination socioéconomique des résultats peu élevés et des proportions d'élèves très peu performants (sous le niveau de compétence 2 du PISA) parmi les moins élevées de notre échantillon.

La Bulgarie présente, quant à elle, un profil diamétralement opposé. Elle combine un niveau de performance moyen parmi les plus faibles de l'échantillon, un très fort degré de détermination socioculturelle des performances scolaires, un différentiel de performance entre élèves parmi les plus forts et une forte proportion d'élèves n'ayant pas atteint le niveau de compétence 2 au PISA très importante.

⁶⁴ Sur la Figure 12 les lignes en pointillés étendus rouges marquent les moyennes de l'échantillon pour le degré d'efficacité et d'inégalités des chances de résultats (en maths), les pointillés simples noirs, les moyennes de l'OCDE pour ces deux dimensions.

Figure 12 : Efficacité et inégalités des chances de résultats, profils croisés des systèmes éducatifs de l'échantillon



Si ce bref descriptif met en exergue des profils contrastés en termes de degrés d'efficacité et d'équité scolaires, il s'agit maintenant d'analyser dans quelle mesure les composantes instrumentales des SABP altèrent, le cas échéant, ces dimensions.

5.2 Résultats des analyses de corrélations

Les analyses de corrélations vont nous permettre d'éclairer tout d'abord les liens statistiques existants entre composantes instrumentales des SABP et indicateurs d'efficacité et d'équité scolaires. Deuxièmement, elles mettent en évidence des relations fortes entre les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon et les variables de contrôle mobilisées dans nos analyses, autrement dit, les facteurs liés aux caractéristiques des systèmes économiques et sociaux et celles des facteurs caractérisant la structure de ces systèmes éducatifs⁶⁵. Finalement, nous montrerons aussi que les composantes instrumentales des SABP entretiennent des relations différenciées avec ces facteurs. Ces résultats sont synthétisés dans le Tableau 19 et sont commentés dans les sous-sections suivantes.

Tableau 19 : Relations (corrélations) entre composantes des SABP, efficacité et équité scolaires, variables de contrôle

	1	2	3	4	5	6	7	8	9	10
1. Efficacité Math (log)	1									
2. Inégalité de résultats Maths (log)	-0,47**	1								
3. Inégalités des chances résultats Maths (log)	-0,23	0,56**	1							
4. Inégalité des acquis de base Maths (log)	-0,96**	0,67**	0,35**	1						
5. Composante (SABP) d'évaluation et régulation selon la performance	-0,16	-0,17	-0,12	0,06	1					
6. Composante (SABP) des incitatifs	-0,26*	0,28*	0,30*	0,28*	,00	1				
7. Composante (SABP) de comparaison des performances des établissements scolaires et d'amélioration interne des performances	-0,04	-0,15	-0,32*	-0,01	,00	,00	1			
8. Dépenses Éducation (z)	0,28*	-0,35**	-0,46**	-0,32*	-0,09	-0,28*	0,42**	1		
9. Variance IESCS (z)	-0,37**	0,41**	0,47**	0,41**	0,25	-0,13	-0,48**	-0,58**	1	
10. Stratification horizontale/verticale (z)	0,07	0,48**	0,60**	0,11	-0,41**	,24	-,35**	-0,44**	0,26*	1
	1	2	3	4	5	6	7	8	9	10

⁶⁵ L'analyse des distributions pour l'ensemble de ces variables est placée en annexe. Du fait des asymétries constatées pour les distributions de certains indicateurs d'efficacité et d'équité, nous avons procédé à une transformation logarithmique des données pour l'ensemble de ces indicateurs. Les autres transformations possibles ne permettant pas de pallier les problèmes de normalité (racine carrée par exemple), nous avons en effet opté pour la transformation log linéaire (Cox & Box, 1964; Howell, 2006; Field, 2009). Les analyses de corrélations, les ANOVA et les régressions multiples sont réalisées sur les indicateurs d'efficacité et d'équité scolaires présentés sous leur forme logarithmique.

5.2.1 Relations (corrélations) entre composantes instrumentales des systèmes d'*accountability* basés sur la performance, efficacité et équité scolaires

Les composantes instrumentales des SABP entretiennent-elles un lien significatif avec les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon? Les résultats des analyses de corrélations, présentés dans les sous-sections suivantes, sont organisés par indicateurs d'efficacité et d'équité scolaires.

5.2.1.1 Efficacité et composantes instrumentales des systèmes d'*accountability* basés sur la performance

En termes d'efficacité scolaire, premièrement, les résultats des analyses de corrélations mettent en exergue le fait que seule la composante instrumentale des SABP se rapportant aux incitatifs articulés à l'évaluation des enseignants est liée significativement au degré d'efficacité des systèmes éducatifs de notre échantillon. Cette relation est négative, d'intensité moyenne (maths : $r = - 0,26$, $p = 0,044$). Il apparaît donc que, tendanciellement, les systèmes éducatifs où cette composante instrumentale est la plus développée sont aussi ceux où les performances moyennes des élèves sont les plus faibles.

Qu'en est-il des relations entre composantes des SABP et équité scolaire ?

5.2.1.2 Inégalité de résultats et composantes instrumentales des systèmes d'*accountability* basés sur la performance

En termes d'inégalité de résultats, on observe que la composante d'évaluation et de régulation selon la performance n'entretient aucune relation significative avec cet indicateur en mathématiques.

La composante instrumentale se rapportant aux incitatifs articulés à l'évaluation des enseignants, entretient quant à elle une relation positive, d'intensité moyenne (maths : $r = 0,28$, $p = 0,035$) ce qui suggère que les systèmes éducatifs où cette composante est de plus forte ampleur sont aussi ceux où le degré d'inégalité de résultats est le plus important.

Finalement, aucune relation n'est constatée avec la composante de comparaison externe des performances des établissements scolaires et d'amélioration de ces dernières par le recours à des dispositifs internes.

5.2.1.3 Inégalités des chances de résultats et composantes instrumentales des systèmes d'*accountability* basés sur la performance

Si la composante d'évaluation et de régulation selon la performance n'entretient pas non plus de lien significatif avec le degré d'inégalité des chances de résultats, celle se rapportant aux incitatifs pour les enseignants est, quant à elle, aussi liée à la détermination socioéconomique des résultats des élèves. Cette relation positive, d'intensité moyenne (maths : $r = 0,30$, $p = 0,023$), implique donc que les systèmes éducatifs où cette composante est la plus développée sont aussi ceux où le degré d'inégalité des chances de résultats est le plus important.

La relation s'inverse avec la composante instrumentale de comparaison des performances des établissements scolaires. La relation est en effet négative (maths : $r = - 0,32$, $p = 0,014$) ce qui signifie qu'à l'inverse, les systèmes éducatifs où cette composante instrumentale est de plus forte ampleur sont ceux où la détermination socioéconomique des performances scolaires est la plus faible.

5.2.1.4 Inégalités des acquis de base et composantes instrumentales des systèmes d'*accountability* basés sur la performance

Enfin, on observe que seule la composante instrumentale se rapportant aux incitatifs articulés à l'évaluation des enseignants entretient un lien significatif, positif et d'intensité moyenne (maths : $r = 0,28$, $p = 0,033$) avec le degré d'inégalité des acquis de base.

Au total, on constate des relations différenciées en fonction des indicateurs mobilisés, des composantes des SABP considérées. Il apparaît que la composante d'évaluation et de régulation selon la performance n'entretient de lien avec aucun indicateur d'efficacité ou d'équité.

La composante des SABP ayant trait aux incitatifs articulés aux évaluations des enseignants est quant à elle liée à la fois au degré d'efficacité, mais aussi au degré d'équité scolaire. Ainsi, il apparaît que le degré d'efficacité scolaire est tendanciellement plus faible dans les systèmes éducatifs où cette composante est de plus forte ampleur, mais aussi que les degrés d'inégalité de résultats, des chances de résultats et des acquis de base y sont aussi plus importants.

Enfin, la composante instrumentale de comparaison entre établissements scolaires et d'amélioration de leurs performances par le biais de dispositifs internes ne semble entretenir de lien significatif qu'avec le degré d'inégalité des chances de résultats. Nos résultats suggèrent dès

lors que les systèmes éducatifs où cette composante est de plus forte intensité sont aussi ceux où la détermination socioéconomique des performances des élèves est la plus faible.

Ces résultats sont synthétisés dans le Tableau 20.

Tableau 20 : Synthèse des relations (corrélations) entre composantes instrumentales des SABP, efficacité et équité scolaires

	Efficacité scolaire	Inégalité de résultats	Inégalité des chances de résultats	Inégalité des acquis de base
Évaluation et régulation selon la performance				
Incitatifs	- maths	+ maths	+ maths	+ maths
comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires			- maths	

5.2.2 Relations (corrélations) entre degrés d'efficacité et d'équité scolaires et variables de contrôle

Qu'en est-il des relations entre efficacité scolaire, équité scolaire, facteurs économiques, sociaux et liés à la structure des systèmes d'éducation ? Puisque plusieurs variables de contrôle sont intégrées à nos analyses de régression multiple, nous décrivons, sur base des analyses de corrélations (Tableau 19), les liens que ces derniers entretiennent avec les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon.

Si l'on s'intéresse premièrement à la part des dépenses publiques investies dans l'éducation, il apparaît que cet indicateur est lié à la fois à des degrés d'efficacité, mais aussi d'équité scolaire plus élevés. La relation est en effet significative et positive avec les performances moyennes des systèmes éducatifs (maths : $r = 0,28$, $p = 0,005$), négative et significative avec l'ensemble des indicateurs d'équité (inégalité de résultats : maths : $r = - 0,35$; $p = 0,006$; inégalité des chances de résultats : maths : $r = - 0,46$, $p = 0,000$; inégalité des acquis de base : maths : $r = - 0,32$, $p = 0,013$). En outre, c'est avec la détermination socioéconomique des résultats que cet indicateur entretient la relation la plus forte. Au total, nos résultats suggèrent que les pays et les régions qui investissent une part plus importante de leur richesse dans l'éducation sont les plus performants et les plus équitables de notre échantillon.

Deuxièmement, il apparaît que la structure sociale des systèmes éducatifs entretient elle aussi une relation significative avec les degrés d'efficacité et d'équité scolaires. Nos résultats

suggèrent en effet que les systèmes éducatifs où la variation du niveau socioéconomique entre élèves est la plus importante sont ceux où, d'une part, le degré d'efficacité est le plus faible (maths : $r = - 0,37$, $p = 0,004$) et où, d'autre part, le degré d'iniquité est le plus élevé, tous indicateurs confondus. Ici, c'est avec la détermination socioéconomique des résultats des élèves que la relation est de plus forte intensité (inégalité de résultats maths : $r = 0,41$, $p = 0,001$; inégalité des chances de résultats maths : $r = 0,47$, $p = 0,000$; inégalité des acquis de base maths : $r = 0,41$, $p = 0,001$).

Enfin, si la stratification horizontale et verticale des systèmes éducatifs ne semble pas liée au degré d'efficacité scolaire ni à la proportion d'élèves très peu performants, elle entretient une relation forte et positive avec les degrés d'inégalité de résultats (maths : $r = 0,48$, $p = 0,000$) et d'inégalité des chances de résultats (maths : $r = 0,60$, $p = 0,000$). La relation de cette variable avec ce dernier indicateur d'équité est, en outre, celle qui présente la plus forte intensité de la série d'analyse.

Pour résumer, il apparaît que les facteurs analysés ici entretiennent des relations significatives et, pour la plupart, de forte intensité avec les indicateurs d'efficacité et d'équité scolaires. Ceci met en avant l'importance de les intégrer à notre analyse pour observer les effets nets des composantes des SABP (et par la suite, des types empiriques de SABP) sur les degrés d'efficacité et d'équité scolaires.

5.2.3 Relations (corrélations) entre composantes instrumentales des systèmes d'*accountability* basés sur la performance et variables de contrôle

Enfin, avant de présenter les résultats des modèles permettant d'appréhender les effets bruts et nets des composantes instrumentales des SABP, nous avons aussi souhaité observer dans quelle mesure les facteurs économiques, sociaux et structurels entretiennent des relations significatives avec les composantes instrumentales des SABP (Tableau 19).

On observe ainsi que la composante instrumentale de comparaison des performances des établissements scolaires et d'amélioration de ces dernières par le recours à des dispositifs internes entretient une relation significative avec l'ensemble des facteurs considérés, positive avec les dépenses publiques en éducation ($r = 0,42$, $p = 0,001$) et négative avec la variation du

niveau socioéconomique des élèves ($r = - 0,48$, $p = 0,000$) ainsi qu'avec la stratification des systèmes scolaires ($r = - 0,35$, $p = 0,006$).

Les composantes instrumentales d'évaluation selon la performance et d'incitatifs ne sont quant à elles corrélées qu'avec l'indicateur de stratification pour la première ($r = -, 041$, $p = 0, 001$) et les dépenses publiques en éducation pour la seconde ($r = - 0,28$, $p = 0,032$).

Si l'on constate des relations différenciées entre les facteurs qui seront mobilisés comme contrôle dans les analyses multivariées et les composantes instrumentales des SABP, ces résultats (voir tableau ci-dessous) seront surtout utiles pour interpréter les effets nets de ces dernières sur les degrés d'efficacité et d'équité scolaires, c'est-à-dire une fois que ces facteurs clefs sont maintenus constants.

5.3 Effets bruts des composantes instrumentales des systèmes d'*accountability* basés sur la performance: résultats des analyses de régression simple

Qu'en est-il des effets bruts des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires? Il s'agit ici d'observer l'effet des composantes instrumentales sur ces dimensions, sans variable de contrôle, au moyen d'analyses de régression simple.

Si les analyses de corrélations ont mis en avant des relations significatives entre certaines composantes des SABP et les indicateurs d'efficacité et d'équité scolaires, nous observons que la part de variation des degrés d'efficacité et d'équité scolaires expliquée par ces dernières est, le cas échéant, relativement faible (Tableau 21).

La composante instrumentale se rapportant aux incitatifs explique environ 5 % du degré d'efficacité scolaire ($F_{(1, 58)} = 4,233$, $p = 0,044$), 6 % du degré d'inégalité de résultats ($F_{(1, 58)} = 4,664$, $p = 0,035$), 7 % du degré d'inégalités des chances ($F_{(1, 58)} = 5,496$, $p = 0,023$) et 6 % du degré d'inégalité des acquis de base ($F_{(1, 58)} = 4,798$, $p = 0,033$).

La composante instrumentale de comparaison des établissements scolaires et d'amélioration de leurs performances serait, quant à elle, liée à une variation d'environ 9 % du degré d'inégalité des chances de résultats ($F_{(1, 58)} = 6,458$, $p = 0,014$).

Aucun effet brut de la composante d'évaluation et de régulation selon la performance ne ressort finalement de nos analyses.

Tableau 21 : Effets bruts des composantes instrumentales des SABP sur les degrés d'efficacité et d'équité scolaires — résultats des analyses de régression simple

	Efficacité (log)	Inégalité résultats (log)	Inégalité chances résultats (log)	Inégalité acquis de base (log)
Évaluation et régulation selon la performance				
R ²	ns	ns	ns	ns
Incitatifs directs pour les enseignants				
R ²	0,053*	0,059*	0,072*	0,061*
Comparaison des performances des établissements scolaires et amélioration des résultats				
R ²	ns	ns	0,086**	ns

** significatif au seuil de 0,01 ; *significatif au seuil de 0,05 ; ns = non significatif

5.4 Effets nets des composantes instrumentales des systèmes d'*accountability* basés sur la performance sur les degrés d'efficacité et d'équité scolaires : résultats des analyses de régressions multiples

Enfin, est-ce que ces effets, faibles au demeurant, se maintiennent lorsque certains facteurs clefs sont pris en compte ? Il s'agit ici d'analyser les relations des composantes instrumentales des SABP avec l'efficacité et l'équité scolaires en intégrant aux modèles les variables de contrôle. Cela nous permet donc d'analyser ces liens dans le cadre d'un raisonnement « toutes choses égales par ailleurs ».

Les résultats des analyses de régressions multiples, où l'effet de chaque composante instrumentale est testé indépendamment des autres composantes, sont synthétisés dans le tableau 22. Les modèles 1 analysent l'effet net de la composante d'évaluation et de régulation selon la performance sur les degrés d'efficacité et d'équité scolaires, les modèles 2 celui de la composante d'incitatifs articulés aux évaluations des enseignants, les modèles 3 celui de la composante de comparaison des établissements scolaires et d'amélioration de leurs performances.

Tous les modèles intègrent les mêmes variables de contrôle pour chaque indicateur d'efficacité et d'équité scolaires : les dépenses publiques en éducation, la variation du niveau socioéconomique des élèves, ainsi que l'indicateur de stratification horizontale et verticale. Toutefois, la variation du niveau socioéconomique des élèves est exclue des modèles où l'inégalité des chances de

résultats est la variable indépendante (le niveau socioéconomique des élèves et son influence sur les performances scolaires est en effet pris en compte dans la construction de l'indicateur d'inégalités des chances de résultats).

Nous mettons d'abord en exergue l'influence relativement stable et importante des variables de contrôle dans nos modèles, pour observer et commenter par la suite si l'effet des composantes instrumentales des SABP (incitatifs et comparaison entre établissements) se maintient « toutes choses égales par ailleurs », c'est à dire lorsque ces facteurs clefs sont intégrés aux analyses.

Puisque, sans grande surprise, à la lumière des analyses précédentes (voir Tableau 21), nous n'observons aucun effet de la composante d'évaluation et de régulation selon la performance, nous ne commentons pas ces modèles, mais nous centrons sur les modèles subséquents.

Tableau 22 : Effets nets des composantes instrumentales sur le degré d'efficacité et d'équité scolaires — résultats des analyses de régressions multiples

	a) Efficacité scolaire (log)	b) Inégalité résultats (log)	c) Inégalité chances de résultats (log)	d) Inégalité acquis de base (log)
MODÈLES 1	β	β	β	β
<i>Évaluation et régulation selon la performance</i>	0,053	- 0,114	0,071	- 0,073
Dépenses éducation (z)	0,201	- 0,006	- 0,226 †	- 0,150
Variance IESCS (z)	- 0,334 *	0,349 *		0,359 *
Stratification horizontale et verticale (z)	0,262	0,343 *	0,526 **	- 0,074
F	3,144	6,676	12,755	3,005
R2	0,189	0,331	0,410	0,182
R2 ajusté	0,129 **	0,281 **	0,378 **	0,121 *
MODÈLES 2				
<i>Incitatifs</i>	- 0,399 **	0,291 *	0,122	0,381 **
Dépenses éducation (z)	0,001	0,153	- 0,224 †	0,045
Variance IESCS (z)	- 0,499 **	0,447 **		0,509 **
Stratification horizontale et verticale (z)	0,289 **	0,366 **	0,469 **	-0,088
F	6,052	8,531	13,278	5,516
R2	0,310	0,387	0,420	0,290
R2 ajusté	0,258 **	0,382 **	0,388 **	0,237 **
MODÈLES 3				
<i>Comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires</i>	- 0,268 †	0,201	-0,053	0,280 †
Dépenses éducation (z)	0,225	-0,012	- 0,233 †	-0,172
Variance IESCS (z)	- 0,413 **	0,386 **		0,435 **
Stratification horizontale et verticale (z)	0,176	0,449 **	0,475 *	0,026
F	4,215	7,291	12,673	4,119
R2	0,238	0,351	0,409	0,234
R2 ajusté	0,181 **	0,303 **	0,376 **	0,177 **

** significatif au seuil de 0,01 %; *significatif au seuil de 0,05 % ; † significatif au seuil de 0,10; quand il n'y a pas d'indication, les résultats ne sont pas significatifs.

5.4.1 Les relations entre variables de contrôle, degrés d'efficacité et d'équité scolaires

Premièrement il nous faut mettre en avant l'influence sensiblement importante des variables de contrôle sur les degrés d'efficacité et d'équité scolaires dans les systèmes éducatifs de notre échantillon.

S'agissant des dépenses publiques en éducation, nos résultats suggèrent une relation significative négative avec les degrés d'inégalités des chances de résultats. La relation est en effet constatée dans tous les modèles bien qu'à des seuils de significativité limites (modèle 1 – c) : $\beta = -0,226$, $p = 0,070$; modèle 2 – c) : $\beta = -0,224$, $p = 0,060$; modèle 3 – c) : $\beta = -0,233$, $p = 0,053$).

La variation du niveau socioéconomique des élèves est quant à elle liée significativement et négativement avec le degré d'efficacité, mais aussi d'équité scolaire. Il ressort premièrement que, dans tous les modèles, cette dernière altère à la baisse le degré d'efficacité scolaire (modèle 1 – a) : $\beta = -0,334$, $p = 0,038$; modèle 2 – a) : $\beta = -0,499$, $p = 0,002$; modèle 3 – a) : $\beta = -0,413$, $p = 0,010$).

Deuxièmement, si l'on s'intéresse aux inégalités de résultats, il apparaît que la variation du niveau socioéconomique des élèves est significativement liée à des différentiels de performance plus importants entre les élèves les plus performants et les plus faibles (modèle 1 – b) : $\beta = 0,349$, $p = 0,018$; modèle 2 – b) : $\beta = 0,447$, $p = 0,000$; modèle 3 – b) : $\beta = 0,386$, $p = 0,009$).

Il en va de même en ce qui a trait aux inégalités des acquis de base. Dans tous les modèles, la relation entre la variation du niveau socioéconomique des élèves et cet indicateur est en effet significative et négative (modèle 1 – d) : $\beta = 0,349$, $p = 0,018$; modèle 2 – d) : $\beta = 0,447$, $p = 0,000$; modèle 3 – d) : $\beta = 0,386$, $p = 0,009$).

La stratification des systèmes scolaires entretient quant à elle une relation systématiquement positive avec le degré d'inégalité de résultats (modèle 1 – b) : $\beta = 0,343$, $p = 0,022$; modèle 2 – b) : $\beta = 0,366$, $p = 0,004$; modèle 3 – b) : $\beta = 0,449$, $p = 0,001$) et avec le degré d'inégalité des chances de résultats (modèle 1 – d) : $\beta = 0,526$, $p = 0,022$; modèle 2 – d) : $\beta = 0,469$, $p = 0,000$; modèle 3 – d) : $\beta = 0,475$, $p = 0,000$). Autrement dit, la stratification des systèmes scolaires est liée à des différences de performances entre les élèves les plus performants et les plus faibles plus élevées, ainsi qu'à un niveau de détermination socioéconomique des performances des élèves plus important.

Enfin, une relation significative entre degré d'efficacité scolaire et stratification des systèmes éducatifs n'apparaît que lorsque la composante se rapportant aux incitatifs articulés aux évaluations des enseignants est intégrée au modèle (modèle 2 – a) : $\beta = 0,289$, $p = 0,027$). La stratification serait donc, dans ce cas, associée à des performances moyennes plus élevées.

5.4.2 Effets nets de la composante instrumentale des incitatifs directs articulés aux évaluations des enseignants : résultats des analyses de régressions multiples

Qu'en est-il des effets des composantes instrumentales des SABP lorsque ces facteurs sont pris en compte? On observe tout d'abord que la relation de la composante instrumentale ayant trait aux incitatifs articulés à l'évaluation des enseignants reste significative, d'une part avec l'indicateur d'efficacité scolaire et, d'autre part, avec les indicateurs d'équité scolaire, à l'exception de l'indicateur d'inégalité de résultats.

En termes d'efficacité scolaire (modèle 2 – a) : $F_{(4, 58)} = 6,052$, $p = 0,00$; $R^2=26\%$), nos résultats suggèrent que l'influence de cette composante sur le degré d'efficacité scolaire reste significative ($\beta = -0,399$, $p = 0,003$) lorsque la part des dépenses en éducation, la variation du milieu socioéconomique des élèves et la stratification des systèmes éducatifs sont prises en compte dans l'analyse. Tel que mis en exergue par les résultats des analyses de corrélation, l'effet est négatif, ce qui implique que cette composante est associée à des performances moyennes plus faibles dans les systèmes éducatifs de notre échantillon.

L'influence de cette composante sur les degrés d'inégalité de résultats (modèle 2 – b) : $F_{(4, 58)} = 8,531$, $p = 0,00$; $R^2=38\%$) et d'inégalité des acquis de base (modèle 2 – d) : $F_{(4, 58)} = 5,516$, $p = 0,00$; $R^2=24\%$) va ici dans le sens d'une altération à la baisse. En d'autres termes, nos résultats suggèrent que la composante instrumentale des incitatifs est associée à des degrés d'inégalité de résultats (modèle 2 – b) : $\beta = 0,291$, $p = 0,020$) et d'inégalité des acquis de base (modèle 2 – d) : $\beta = 0,381$, $p = 0,005$) plus importants, même si l'on contrôle pour les dépenses publiques en éducation, la variation du milieu socioéconomique des élèves ou encore l'organisation stratifiée des parcours scolaires.

Par contre, l'effet de cette composante sur le degré d'inégalité des chances de résultats, d'abord constaté dans les analyses de régression simple, ne se maintient pas lorsque ces facteurs sont pris en compte dans les analyses (modèle 2 – c) : $\beta = 0,122$, $p =$ non significatif).

5.4.3 Effets nets de la composante de comparaison externe et dispositifs internes d'amélioration des performances des établissements scolaires : résultats des analyses de régressions multiples

Il apparaît par ailleurs que la composante de comparaison des performances des établissements scolaires entretient une relation négative avec les degrés d'efficacité scolaire et d'inégalité des acquis de base, bien que les seuils de significativité soient ici limites.

Le modèle (modèle 3 – a) : $F_{(4, 58)} = 4,215$, $p = 0,005$; $R^2=18\%$) étudiant cette dimension suggère en effet que cette composante est liée à des degrés d'efficacité scolaires plus faibles, mais à un seuil de significativité supérieur à 0,5 ($\beta = - 0,268$, $p = 0,063$). D'autre part, elle est (modèle 3 – d) : $F_{(4, 58)} = 4,119$, $p = 0,006$; $R^2=18\%$) liée significativement à des inégalités des acquis de base plus importantes (que les autres composantes?) ($\beta = 0,280$, $p = 0,053$). Le seuil de significativité est ici proche de 0,5.

Il faut cependant noter que les liens entre cette composante, efficacité et inégalités des acquis de base n'ont été mis en exergue ni par les analyses de corrélations, ni par les analyses de régression simple. Ils n'apparaissent qu'après intégration des variables de contrôle dans les modèles d'analyse. Nous avons par ailleurs constaté que les corrélations entre cette composante et les variables de contrôle étaient relativement fortes, plus particulièrement avec la variation du niveau socioéconomique des élèves. On peut donc se demander si cette relation entre composante de comparaison des performances des établissements scolaires et les dimensions étudiées ci-dessus n'est pas le produit d'un artéfact. Le coefficient de détermination du modèle étant en outre très faible, ce qui témoigne de sa mauvaise qualité d'ajustement. Ce résultat est donc soumis à caution.

Par ailleurs, alors que les analyses préliminaires avaient, à l'inverse, mis en exergue une relation avec l'indicateur d'inégalité des chances de résultats, cette dernière n'est plus significative après l'introduction des variables de contrôle. En d'autres termes, la composante instrumentale de comparaison des performances des établissements scolaires et d'amélioration de ces dernières par des dispositifs internes ne semble entretenir aucun lien avec le degré d'inégalité des chances de résultats, quand la part des dépenses publiques en éducation, la variation du niveau socioéconomique des élèves et la stratification du système sont prises en compte.

5.5 Synthèse

Que retenir de cette première série d'analyses ?

On constate premièrement que seule la composante des incitatifs semble entretenir des relations significatives avec l'efficacité et l'équité scolaires, avec toutefois des variations en fonction des conceptions de l'équité analysées.

Les relations entre cette composante instrumentale, degré d'efficacité et d'équité scolaires, ont été tout d'abord mises en exergue par l'ensemble des analyses préliminaires. La relation mise en évidence par les corrélations est d'intensité faible et négative avec le degré d'efficacité, positive et d'intensité faible avec le degré d'équité scolaire (inégalités de résultats, des chances de résultats, des acquis de base). Ceci implique donc que les systèmes éducatifs où cette composante de l'*accountability* est de plus forte intensité sont aussi ceux qui, tendanciellement, sont les moins efficaces et les moins équitables. La proportion de la variation de ces dimensions expliquées par cette composante oscille en outre entre 5 % et 7 %, ce qui est très faible.

Ces relations restent significatives lorsque les variables de contrôle sont intégrées aux modèles, excepté pour l'indicateur d'inégalité des chances de résultats. Nos analyses suggèrent donc que les incitatifs entretiennent des liens significatifs avec l'efficacité scolaire, les inégalités de résultats et des acquis de base au-delà des dépenses publiques en éducation, de la variation du niveau socioéconomique des élèves et de la stratification du système. Ceci n'est toutefois pas le cas lorsque la détermination socioéconomique des performances scolaires est examinée/analysée. Autrement dit, nos analyses suggèrent que les incitatifs influenceraient à la baisse l'efficacité scolaire dans les systèmes éducatifs de notre échantillon, de même que l'égalité de résultats et des acquis de base.

Si la composante instrumentale se rapportant à la comparaison externe/amélioration interne des performances des établissements scolaires entretient, d'après les analyses préliminaires, une relation significative négative avec le degré d'inégalité des chances de résultats dont elle explique environ 9 % de la variation, la significativité de cette dernière relation n'est pas maintenue après contrôle. Autrement dit, lorsque les dépenses publiques en éducation, la variation du milieu socioéconomique des élèves et la stratification des systèmes sont prises en compte, cette composante ne semble pas produire d'effets significatifs sur l'inégalité des chances de résultats.

La difficulté posée, entre autres, par l'analyse des composantes instrumentales des SABP à titre indépendant relève de la comparabilité de nos résultats avec ceux de la littérature. Nous pourrions rapprocher l'impact négatif des incitatifs sur l'efficacité et l'équité scolaires (égalité de résultats et égalité des acquis de base) des travaux qui portent sur les SABP à forts enjeux, les enjeux pour les acteurs constituant en effet l'un des piliers centraux de ces systèmes d'*accountability* (Carnoy & Loeb, 2002 ; Harris & Herrington, 2006 ; Dupriez & Mons, 2011). Comme nous l'avons vu dans la revue de littérature réalisée en amont, il émerge tendanciellement des études qui s'intéressent aux effets des SABP à forts enjeux que des conséquences matérielles et sociales importantes pour les acteurs auraient des effets négatifs sur l'équité scolaire, principalement du fait des comportements stratégiques mis en œuvre par ces derniers placés sous la menace de ces dispositifs (Figlio & Loeb, 2011; Jacob, 2001, 2005; Mons, 2009; Webb, 2005). Cependant notre composante instrumentale réfère à une catégorie spécifique d'enjeux, plus restreinte que ceux analysés par ces études, à savoir les conséquences matérielles et sociales qui font suite aux évaluations des enseignants.

L'étude de Wößmann (Wößmann, 2011) est ici intéressante puisqu'elle s'intéresse plus spécifiquement aux incitatifs visant les enseignants et à leurs effets sur l'efficacité scolaire. Il interroge en effet l'impact des primes à la performance pour les enseignants sur les résultats des élèves. La recherche, menée sur un échantillon restreint de 28 pays, à partir des données recueillies dans le cadre de l'*International Indicators of Educational Systems Survey* de l'OCDE et de celles du PISA 2003, conclut que ces dispositifs auraient un impact positif sur le degré d'efficacité scolaire. Autrement dit, l'octroi de primes à la performance aux enseignants engagerait une augmentation des performances moyennes des élèves. Si ces résultats contrastent avec les nôtres, on peut faire l'hypothèse que les divergences de conclusions sont en partie dues au fait que notre composante instrumentale ne se focalise par uniquement sur les incitatifs financiers (bien que ces derniers apportent une contribution majeure à la construction de cette dernière), mais prend en compte des incitatifs de différentes natures au moyen d'un indicateur composite. Par ailleurs, la question des effets des primes à la performance sur l'équité scolaire n'est pas examinée dans l'étude de Wößmann. D'autres études sur la question seraient donc nécessaires pour pouvoir mettre en discussion et comparer les résultats de notre recherche.

Enfin, il faut noter l'influence relativement stable des variables de contrôle sur l'efficacité et l'équité des systèmes scolaires. Les dépenses publiques en éducation sont en effet liées tendanciellement à des degrés d'inégalité des chances de résultats plus faibles, résultats déjà mis en exergue par la littérature. Toutefois ici, tel que nous l'avons vu, cette relation apparaît à des seuils de significativité limites dans notre étude. On peut donc imaginer, tel que la littérature le suggère, que les facteurs économiques (dépenses publiques en éducation, richesses des pays [PIB]) jouent effectivement dans l'amélioration de l'efficacité scolaire et la réduction des inégalités, mais qu'au-delà d'un certain seuil (développement économique élevé) l'influence de ces derniers devient réduite (Mons, 2009).

La variation du niveau socioéconomique des élèves est, quant à elle, liée à une efficacité scolaire plus faible et à des inégalités plus importantes. Ce constat relativement classique dans la littérature a donc trait à l'influence sensiblement importante du niveau socioéconomique des élèves sur leurs performances scolaires (Marks, 2005; Mons, 2009; OCDE, 2013a), et donc au lien fort, que nombre de politiques éducatives tentent d'atténuer (nous y reviendrons) entre inégalités sociales et inégalités scolaires.

Enfin, la stratification des systèmes éducatifs, autrement nommée différenciation verticale et horizontale, serait quant à elle liée à des inégalités de performances et des chances de résultats plus importantes. Ce constat est en effet appuyé par nombre de recherches sur la question (Dupriez & Dumay, 2006; Mons, 2007; 2008; OCDE, 2013). Fait surprenant toutefois, nos résultats suggèrent que la stratification des systèmes scolaire entretiendrait une relation positive avec le degré d'efficacité scolaire. Néanmoins, cette relation n'apparaît dans nos analyses que lorsque la composante d'*accountability* ayant trait aux incitatifs est prise en compte dans les modèles. Cette relation entre efficacité scolaire et stratification des systèmes éducatifs contredit en partie les résultats des dernières études sur la question, mais tend à aller dans le sens du présumé d'efficacité attaché à la création de filières dans les systèmes éducatifs, pourtant remis en question par certaines études empiriques (Mons, 2004). Il faut préciser ici que notre indicateur prend en compte à la fois le nombre de filières rencontrées à 15 ans, la part des scores des élèves attribuable à l'appartenance à un établissement scolaire et la pratique du redoublement, agrégés dans la construction de notre indicateur.

À la lumière de nos résultats, une piste interprétative possible serait que la sélection opérée au sein des systèmes éducatifs favoriserait les élèves les plus performants, leurs performances impactant dès lors à la hausse la moyenne des scores, tout en pénalisant les plus faibles puisque les inégalités de résultats, des chances de résultats sont affectés à la baisse par cette dernière. Néanmoins, puisque la relation entre stratification du système et efficacité n'apparaît que lorsque les incitatifs articulés aux évaluations des enseignants sont intégrés à l'analyse, cette dernière semble peu stable. Il s'agirait de creuser davantage cette question dans des recherches subséquentes.

Au total, nos résultats suggèrent que les composantes instrumentales des SABP auraient un effet différencié sur les degrés d'efficacité et d'équité scolaires des systèmes d'éducation de notre échantillon. Cet effet, négatif le cas échéant, ne s'observe toutefois que pour une seule composante instrumentale (ayant trait à la mise en place d'incitatifs direct pour les enseignants) et restent faibles. Les analyses de régression simple suggèrent en effet que la composante d'incitatifs directs pour les enseignants explique environ 5 % de la variation des degrés d'efficacité scolaire, 6 % du degré d'inégalité de résultats et des acquis de base. Si cet effet se maintient lorsque les dépenses publiques en éducation, la variation du niveau socioéconomique des élèves et la stratification du système scolaire sont intégrées à l'analyse, ce n'est en outre pas le cas lorsque le degré d'inégalité des chances de résultats est interrogé (aucun effet pour la composante d'incitatifs n'est constatée). En outre, il faut observer que l'intégration des variables de contrôle aux modèles engage une augmentation importante des coefficients de détermination de ces derniers (par rapport aux modèles de régression simple). Cette augmentation est d'environ 20 % si l'on s'intéresse à l'efficacité scolaire (modèle 2 a), de plus de 30 % en ce qui a trait aux inégalités de résultats et des chances de résultats (modèle 2 b et c), d'environ 18 % s'agissant des inégalités des acquis de base. Autrement dit, ces facteurs semblent jouer un rôle davantage déterminant dans la variation des degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon.

SECTION 2 : Types de systèmes d'*accountability* basés sur la performance, degrés d'efficacité et d'équité scolaires, des relations différenciées?

Qu'en est-il des effets des types empiriques de SABP sur les degrés d'efficacité et d'équité scolaires? Observe-t-on des relations différenciées en fonction des types empiriques de SABP sur l'efficacité et l'équité scolaires? Les sous-sections suivantes permettent d'apporter des éléments de réponse à cette question. Elles examinent les relations des types empiriques de SABP avec les degrés d'efficacité et d'équité scolaires.

Pour rappel, notre typologie empirique exploratoire a permis de caractériser quatre types de SABP :

- 1) une *accountability* de faible intensité peu instrumentée (type 1);
- 2) une *accountability* axée sur l'évaluation de la performance (type 2);
- 3) une *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance (type 3);
- 4) une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants (type 4).

Dans la même logique que dans les sous-sections précédentes, nous avons tout d'abord réalisé des analyses préliminaires permettant d'étudier le lien entre types empiriques de SABP, efficacité et équité scolaires. Dans les sous-sections suivantes nous présentons donc premièrement un descriptif des degrés d'efficacité et d'équité scolaires par types empiriques de SABP. Les résultats des ANOVA⁶⁶ sont par la suite présentés. Ils nous permettent d'observer la présence de différences statistiquement significatives dans les degrés d'efficacité et d'équité scolaires des types de SABP et donc, de valider, le cas échéant, les tendances mises en exergue par les statistiques descriptives. Finalement, les analyses de régressions multiples nous permettent d'observer si les relations mises en exergue dans cette première série d'analyses se maintiennent lors d'un raisonnement « toutes choses égales par ailleurs », c'est-à-dire lorsque les variables de contrôle (dépenses publiques en éducation, variation du niveau socioéconomique des élèves, stratification des systèmes scolaires) sont intégrés aux analyses.

⁶⁶ Analyses de variance à un critère de classification où les variables dépendantes (Y) sont les degrés d'efficacité et d'équité scolaires (égalité de résultats, des chances de résultats et des acquis de base), le critère de classification étant les types de SABP.

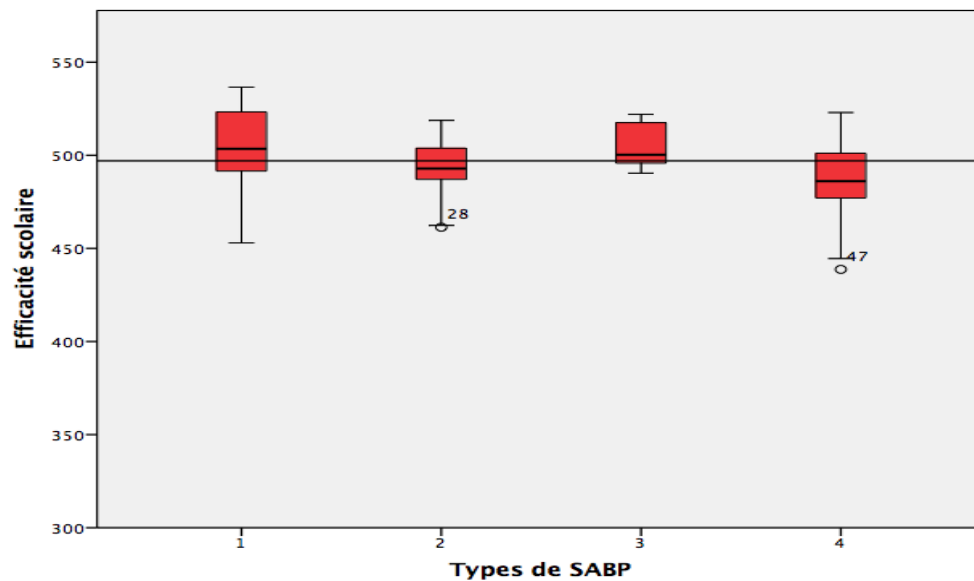
5.6 Descriptif des degrés d'efficacité et d'équité scolaires par types de systèmes d'accountability basés sur la performance

Qu'en est-il des caractéristiques des types de SABP en termes d'efficacité scolaire et d'équité scolaire?

Si l'on s'intéresse premièrement au degré d'efficacité scolaire, la Figure 13 illustre les distributions des types de SABP pour cette dimension. La ligne horizontale sur la figure rend compte de la moyenne de l'échantillon en termes d'efficacité scolaire, les lignes horizontales au sein de chaque boîte représentent la médiane pour chaque type de SABP.

On observe premièrement que les distributions présentent, pour chaque type de SABP, une asymétrie négative. Les médianes sont par ailleurs plus basses pour les types 2 et 4 de SABP, ce qui indique que ces types de SABP seraient tendanciellement moins efficaces que les types 1 et 3. En outre, on observe la présence de valeurs extrêmes pour les types 2 et 4, ce qui indique des degrés d'équité scolaire particulièrement faibles pour certains systèmes éducatifs caractérisés par ces types empiriques de SABP (la Grèce pour le type 1 ; la Bulgarie pour le type 4).

Figure 13 : Efficacité scolaire et types de SABP – examen des distributions



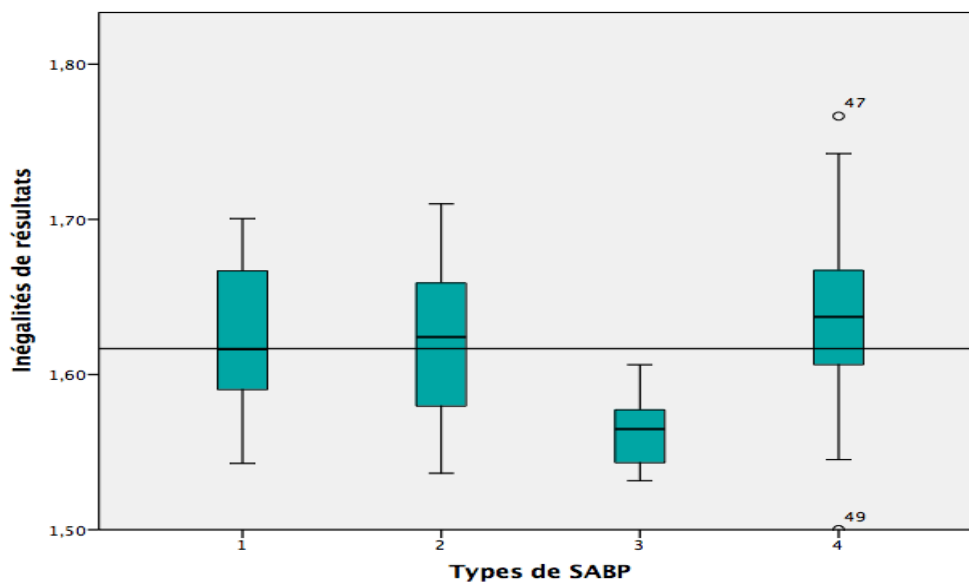
En termes d'inégalité de résultats, la Figure 14 met en exergue le fait que la distribution du type 3, qui présente une légère asymétrie positive, est par ailleurs plus homogène et moins étendue que les autres types des SABP. En outre, la médiane pour ce type est inférieure à celle des autres types de SABP, ce qui indique que le degré d'inégalités de résultats y est sensiblement plus faible.

Il faut par ailleurs noter que les distributions des types 1, 2 et 4 sont clairement plus étendues que celle du type 3, impliquant de plus fortes variations en termes de degrés d'égalité de résultats entre les systèmes éducatifs qui y sont regroupés au sein d'un même type empirique.

Nous observons en outre une légère augmentation des médianes entre les types 1, 2 et 4, ce dernier présentant clairement la médiane la plus élevée. Cela suggère que les degrés d'inégalités de résultats augmenteraient légèrement et progressivement du type 1 au type 4, le type 3 de SABP ne s'inscrivant par ailleurs pas dans cette tendance.

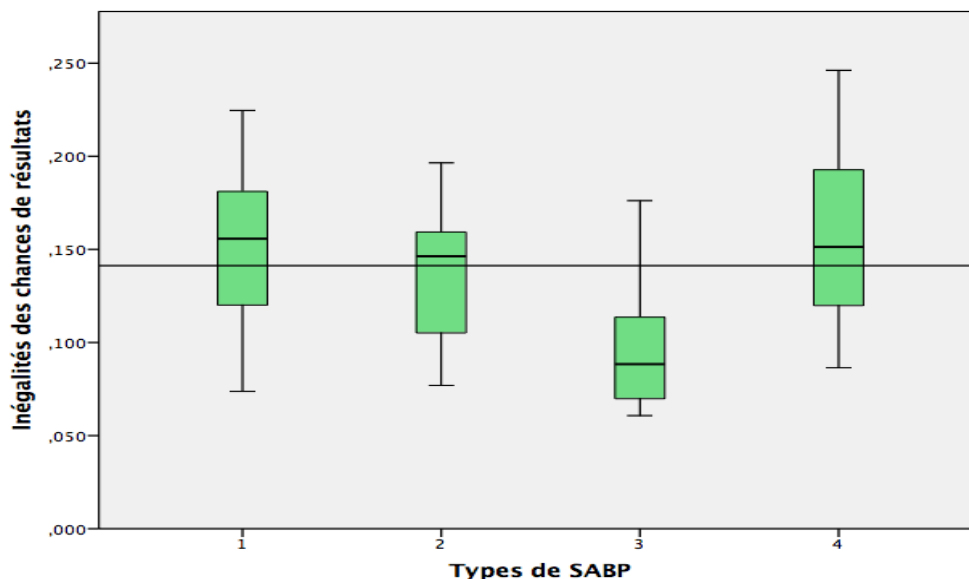
Finalement, il faut aussi observer la présence de valeurs extrêmes au sein du type 4 de SABP, ce qui implique que certains systèmes éducatifs s'y distinguent des autres par un degré d'inégalités de résultats sensiblement plus faible (en l'occurrence l'Estonie M=1,50) ou nettement plus élevé (comme la Bulgarie M=1,77).

Figure 14 : Inégalités de résultats et types de SABP – examen des distributions



S'agissant du degré d'inégalités des chances de résultats, si la Figure 15 permet d'observer une diminution progressive des médianes entre les types 1, 2 et 3, elle met par ailleurs en relief que la médiane du type 3 est ici encore sensiblement plus faible que celle des autres types de SABP. Ceci suggère que ce type de SABP semble aussi se caractériser par un degré d'inégalité des chances de résultats plus faible que les autres types.

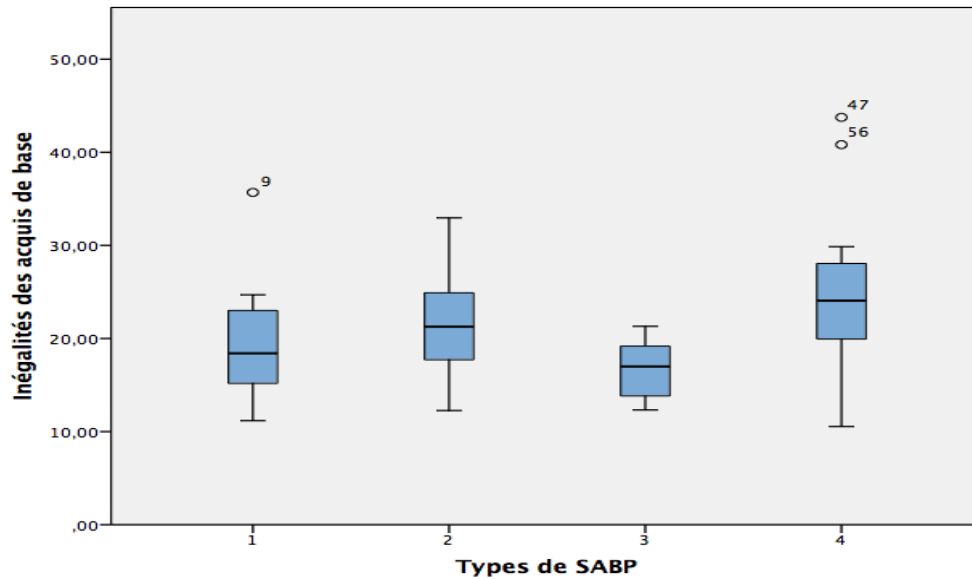
Figure 15 : Inégalités des chances de résultats et types de SABP – examen des distributions



Finalement, en termes d'inégalité des acquis de base, la Figure 16 permet d'observer que le type 4 présente la médiane la plus élevée et le type 3 enregistre encore la médiane la plus basse. Autrement dit, il semblerait que le degré d'inégalités des chances de résultats soit plus important dans le type 4 que dans les autres types de SABP et, à l'inverse, plus faible au sein du type 3. On observe en outre que les distributions des types 1 et 3 sont moins étendues que celle des types 2 et 4, la distribution de ce dernier présentant une forte asymétrie positive.

Notons par ailleurs la présence de valeurs extrêmes au sein des types 1 et 4 où certains systèmes éducatifs (la Grèce au sein du type 1 [M=35,69]; la Bulgarie [M=43,76] et la Roumanie [M=40,82] au sein du type 2) se distinguent clairement par des proportions d'élèves n'ayant pas atteint un niveau de compétence de base (niveau 2 aux tests du PISA) sensiblement plus élevées que les autres systèmes éducatifs appartenant au même type de SABP.

Figure 16 : Inégalités des acquis de base et types de SABP – examen des distributions



Au total, ce descriptif permet de mettre en exergue différentes tendances. Il apparaît tout d’abord que le type 1 de SABP, caractérisé par une *accountability* de faible intensité et peu instrumentée, semble se distinguer par un degré d’efficacité scolaire plus élevé que celui des autres types de SABP.

Le type 4, défini par une *accountability* de forte intensité prévoyant des incitatifs pour les enseignants, se distingue à l’inverse, par un degré d’efficacité scolaire plus faible que les autres types de SABP. Ce dernier se caractérise par ailleurs par des degrés d’inégalités de résultats et des acquis de base tendanciellement plus élevés que les autres types de SABP.

Le type 3 qui se caractérise par une *accountability* axée sur l’évaluation selon la performance et la comparaison externe/amélioration interne de la performance, paraît *a priori* se distinguer par des degrés d’inégalités de résultats et des chances de résultats plus faibles que les autres types de SABP.

Il semble par ailleurs difficile de faire émerger une tendance claire en ce qui concerne le type 2 de SABP, caractérisé par une *accountability* axée principalement sur l’évaluation de la performance. Il se distingue toutefois par un degré d’efficacité scolaire parmi les plus faibles des types empiriques, mais aussi par des degrés d’inégalités de résultats et des acquis de base parmi les plus élevés des types de SABP.

Les tendances suggérées par ces statistiques descriptives sont-elles confirmées par les analyses de variance? Autrement dit existe-il des différences statistiquement significatives de degrés d'efficacité et d'équité scolaire selon les types de SABP ?

5.7 Résultats des ANOVA

Les analyses de variance à un critère de classification nous permettent justement d'observer si les différences des degrés d'efficacité et d'équité scolaires entre les types de SABP suggérées par les analyses descriptives sont effectivement statistiquement significatives. Pour information, les statistiques descriptives (moyenne, écart-type) pour les indicateurs d'efficacité et d'équité scolaires par type de SABP sont placées en annexe, au même titre que les résultats des ANOVA (ANNEXE 7 et ANNEXE 8).

En termes d'efficacité scolaire premièrement, les résultats des ANOVA confirment que les types de SABP influenceraient différemment les degrés d'efficacité scolaire ($F_{(3,55)} = 2,814$, $p = 0,048$). Les résultats du test de Tukey suggèrent que des différences significatives ne s'observent toutefois qu'entre le type 1 ($M = 504,75$; $ET = 21,16$) et le type 4 ($M = 486,82$; $ET = 25,45$). Le degré moyen d'efficacité scolaire du type 1 est donc significativement supérieur à celui du type 4, bien que cette différence ne s'observe qu'à un seuil de significativité de 0,10. Aucune différence statistiquement significative n'est par ailleurs constatée avec les types 2 ($M = 492,70$; $ET = 17,06$) et 3 ($M = 504,95$; $ET = 12,35$).

S'agissant du degré d'inégalités de résultats, les résultats des ANOVA ($F_{(3,55)} = 3,227$, $p = 0,029$) mettent aussi en exergue que les types de SABP influenceraient cette dimension. D'après les résultats du test de Tukey, le type 3 ($M = 1,56$; $ET = 0,02$) se démarque ici des types 1 et 4 par des degrés d'inégalités de résultats plus faibles (type 1 : $M = 1,62$; $ET = 0,05$; type 4 : $M = 1,64$; $ET = 0,07$), ce qui confirme les tendances qui émergent du descriptif réalisé en amont. Il est toutefois à noter qu'aucune différence avec le type 2 (type 2 : $M = 1,62$; $ET = 0,03$), et donc entre le type 1 et 2 de SABP, n'est mise de l'avant par nos analyses.

En termes d'inégalités des chances de résultats, là encore, les résultats des ANOVA ($F_{(3,55)} = 5,546$, $p = 0,002$) et du test de Tukey confirment par ailleurs les tendances descriptives discutées en amont : c'est ici le type 3 ($M = 0,10$; $ET = 0,04$) qui se distingue par un degré d'inégalité des chances de résultats plus faible que les autres types de SABP (type 1 : $M = 0,15$;

ET=0,04 ; type 2 : M=0,14 ; ET=0,03 ; type 4 : M=0,16 ; ET=0,05).

Enfin, s'agissant du degré d'inégalité des acquis de base, le test ANOVA n'est significatif qu'à un seuil relativement plus élevé bien que proche de 0,05 ($F_{(3,55)}=2,672$ $p=0,056$). Selon le test de Tukey, par ailleurs, le type 3 (M=16,71; ET=3,21) se distingue du type 4 (M=24,71; ET=9,39) par un degré moyen d'inégalité des acquis de base plus faible. Le premier (type 3) se caractérise en effet par le degré moyen d'inégalités des acquis de base le plus faible, le second (le type 4) par le degré d'inégalités des acquis de base le plus importantes des types empiriques. Aucune différence n'est toutefois constatée avec les types 1 (M=19,13; ET=5,76) et 2 (M=21,29; ET=5,85) de SABP.

En guise de synthèse, nous reportons les résultats des ANOVA dans le Tableau 23.

Tableau 23 : Types empiriques de SABP, efficacité et équité scolaires : synthèse des résultats des ANOVA

Indicateurs	Différences significatives entre types empiriques de SABP
Efficacité scolaire	Type 1 vs type 4†
Inégalité de résultats	Type 3 vs type 1 et 4*
Inégalité des chances de résultats	Type 3 vs type 1, 2 et 4*
Inégalité des acquis de base	Type 3 vs type 4†

* Les différences de moyennes entre types empiriques sont significatives au seuil de 0,05; † Les différences de moyennes entre types empiriques sont significatives au seuil de 0,10.

Au total, si des différences dans les degrés d'efficacité et d'équité scolaires des types de SABP sont suggérées par ces analyses préliminaires, ces effets potentiellement différenciés se maintiennent-ils une fois que d'autres facteurs clefs sont pris en compte dans les analyses? Autrement dit l'effet différencié des types de SABP sur les degrés d'efficacité et d'équité scolaire se maintient-il « toutes choses égales par ailleurs » ?

5.8 Effets nets des types de systèmes d'*accountability* basés sur la performance sur les degrés d'efficacité et d'équité scolaires : résultats des analyses de régressions multiples

Cette dernière sous-section nous permet de répondre à cette question et d'appréhender les effets nets des types de SABP sur l'efficacité et l'équité scolaires. Nous y présentons les résultats des analyses de régressions multiples qui intègrent les dépenses publiques en éducation, la variation du niveau socioéconomique des élèves ainsi que la stratification des systèmes scolaires comme variables de contrôle.

Pour rappel, nous avons choisi ici une méthode de régression hiérarchique par blocs⁶⁷ pour tester les liens entre types de SABP, efficacité et équité scolaires. Le Tableau 24, qui présente les résultats des analyses de régressions multiples, rend compte premièrement des effets bruts des types de SABP sur l'efficacité et l'équité scolaires, c'est-à-dire sans prise en compte des variables de contrôle (effets bruts-modèles 1). Les effets nets y sont aussi présentés (effets nets-modèles 2). Cette méthode de régression nous permet ainsi d'appréhender d'une part la variation de la proportion des degrés d'efficacité et d'équité scolaires expliquée par les modèles lorsque, tout d'abord, seuls les types de SABP sont pris en compte, puis lorsque d'autres facteurs clefs sont intégrés aux analyses. D'autre part, ceci nous permet d'observer si les effets bruts des types de SABP, le cas échéant, se maintiennent lorsque d'autres facteurs clefs sont pris en compte.

Rappelons par ailleurs que la classe de référence, le type 1 où l'*accountability* est de plus faible intensité, est exclu des modèles. L'interprétation des résultats se fait donc en rapport à cette classe de référence.

⁶⁷ Voir chapitre méthodologie, sous-section Analyses de régressions multiple.

Tableau 24 : Effets des types empiriques de SABP sur les degrés d'efficacité et d'équité scolaires

	a) Efficacité	b) Inégalités de résultats	c) Inégalité des chances de résultats	d) Inégalité des acquis de base
EFFETS BRUTS — MODÈLES 1				
	β	β	β	β
Type 2	- 0,253† (0,085)	- 0,060	-0,130	0,160
Type 3	0,008	- 0,365**	-0,478**	- 0,123
Type 4	- 0,364**	0,083	0,051	0,308*
F	2,814	3,227	5,546	2,672
R2	0,133	0,150	0,232	0,127
R2 ajusté	0,086*	0,103*	0,190*	0,080†
EFFETS NETS — MODÈLES 2				
Type 2	0,149	- 0,142	- 0,072	-0,167
Type 3	0,098	- 0,125	- 0,210	-0,109
Type 4	- 0,376**	0,172	0,018	0,349**
Dépenses éducation (z)	0,028	0,089	- 0,242†	0,013
Variance IESCS (z)	- 0,577**	0,465**		0,591**
Stratification horizontale et verticale (z)	0,363*	0,288†	0,375**	-0,172
F	4,584	5,258	8,253	4,200
R2	0,346	0,378	0,438	0,326
R2 ajusté	0,270**	0,306**	0,385**	0,249**

** significatif au seuil de 0,01 %; *significatif au seuil de 0,05 %; † significatif au seuil de 0,10 %; quand il n'y a pas d'indication, les résultats ne sont pas significatifs.

5.8.1 Types de systèmes d'*accountability* basés sur la performance et efficacité scolaire

Qu'en est-il des effets des types de SABP sur l'efficacité scolaire? En cohérence avec les ANOVA, le premier modèle sans contrôle (modèle 1 – a) : $F_{(6, 58)} = 2,814$, $p = 0,048$; $R^2=9\%$) met en exergue une relation significative entre le degré d'efficacité scolaire et le type 4 caractérisé par une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants (modèle 1 – a) : type 4 : $\beta = - 0,364$, $p = 0,014$). Pour rappel, ce type regroupe la plupart des systèmes éducatifs de l'est de l'Europe, avec l'Angleterre, le Pays de Galles et l'Irlande du Nord, la Suède et les Pays-Bas. Le coefficient de détermination du modèle est toutefois relativement faible. Ce dernier explique en effet moins de 10 % de la variation du degré d'efficacité scolaire dans les systèmes éducatifs de notre échantillon.

L'effet de ce type empirique, au demeurant négatif, reste significatif (modèle 2 – a) type 4 : $\beta = - 0,577$, $p = 0,014$) après intégration des variables de contrôle. L'introduction de ces dernières

implique une augmentation du coefficient de détermination du modèle sensiblement importante – environ 15% (modèle 2 – a) : $F_{(6, 58)} = 4,584$, $p = 0,001$; $R^2=27\%$). Le modèle final explique environ un quart de la variation du degré d'efficacité scolaire. Nos résultats suggèrent donc que ce type empirique de SABP (le type 4) est lié significativement à des degrés d'efficacité scolaires plus faibles que le type de référence où l'*accountability* est de faible intensité et peu instrumentée.

5.8.2 Types de systèmes d'*accountability* basés sur la performance et degrés d'inégalités de résultats

Qu'en est-il des effets des types empiriques de SABP sur le degré d'inégalités de résultats?

Dans la même lignée que les ANOVA, les résultats des modèles de régression sans contrôle (modèle 1 – b) : $F_{(6, 58)} = 3,227$, $p = 0,029$; $R^2=10\%$) mettent en exergue le fait que le type 3 se distingue de la classe de référence par un degré d'inégalités de résultats plus faible. Tel que nous l'avons vu, le degré d'inégalités de résultats est, dans le type 3 en effet, sensiblement plus faible qu'au sein du type 1 (classe de référence).

La relation constatée devient cependant non significative une fois que les variables de contrôle sont intégrées au modèle (modèle 2 – b) : $F_{(6, 58)} = 4,584$, $p = 0,001$; $R^2=30\%$). Seuls les coefficients de la variation du niveau socioéconomique des élèves (modèle 2 – b) : $\beta = -0,465$, $p = 0,006$) et de la stratification horizontale et verticale des systèmes scolaires (modèle 2 – b) : $\beta = -0,288$, $p = 0,066$) sont significatifs. En d'autres termes l'effet (brut) du type 3 ne se maintient pas « toutes choses égales par ailleurs ».

5.8.3 Types de systèmes d'*accountability* basés sur la performance et degrés d'inégalités des chances de résultats

Si l'on s'intéresse maintenant au degré d'inégalité des chances de résultats (modèle 1 — c) : $F_{(6, 58)} = 5,546$, $p = 0,002$; $R^2=19\%$) on observe que le type 3 semble, là encore, se distinguer de la classe de référence par un degré d'inégalité des chances de résultats plus faible (modèle 1 — c) : type 3 : $\beta = -0,478$, $p = 0,006$).

L'effet (effet brut) d'abord constaté entre ce type empirique et le degré d'inégalités des chances de résultats s'annule toutefois lorsque les variables de contrôle sont ajoutées au modèle (modèle

2 - c) : $F_{(6, 58)} = 8,253$, $p = 0,000$; $R^2=39\%$). L'introduction des dépenses publiques en éducation (modèle 2 - c) : $\beta = -0,242$, $p = 0,060$) et de l'indicateur de stratification horizontale et verticale (modèle 2 - c) : $\beta = 0,375$, $p = 0,011$) implique en effet que la relation entre ce type de SABP et l'indicateur d'inégalités des chances de résultats devient non significative. En outre, l'introduction de ces variables de contrôle ajoute une forte contribution au coefficient de détermination du modèle : environ 20%. Le modèle explique alors près de 40 % de la variation du degré d'inégalité des chances de résultats des systèmes éducatifs de notre échantillon.

5.8.4 Types de systèmes d'*accountability* basés sur la performance et degrés d'inégalités des acquis de bases

Enfin, qu'en est-il de la relation entre types de SABP et degré d'inégalité des acquis de base?

Dans le même sens que les résultats des ANOVA, les analyses de régressions multiples suggèrent ici que le type 4 se distingue de la classe de référence par un degré d'inégalité des acquis de base plus élevé, à un seuil de significativité supérieur à 0,05 toutefois (modèle 1 - d) : $F_{(6, 58)} = 2,672$, $p = 0,056$; $R^2=8\%$). La relation négative constatée dans le modèle sans variables contrôle (modèle 1 - d) : type 4 : $\beta = 0,308$, $p = 0,036$) se maintient par ailleurs lorsque l'ensemble des variables sont intégrées aux analyses.

Les résultats de ce dernier modèle (modèle 2 - d) : $F_{(6, 58)} = 4,200$, $p = 0,002$; $R^2 = 25\%$) suggèrent donc que le type 4 de SABP est lié à des proportions d'élèves n'ayant pas atteint un niveau de compétences de base à 15 ans plus élevées (modèle 2 - d) : type 4 : $\beta = 0,349$, $p = 0,017$). Il faut par ailleurs noter que, dans ce dernier modèle, seule la variation du niveau socioéconomique des élèves est aussi liée à des degrés d'inégalités des acquis de base plus élevés (modèle 2 - d) : variation du niveau socioéconomique des élèves : $\beta = 0,591$, $p = 0,001$), les coefficients standardisés des dépenses publiques en éducation et de la stratification des systèmes d'éducation n'étant pas significatifs dans ce modèle.

5.9 Synthèse

Que retenir de cette deuxième série d'analyse ?

Les résultats de cette dernière série d'analyses ont permis tout d'abord de mettre en exergue des différences dans les degrés moyens d'efficacité et d'équité scolaires en fonction des types empiriques de SABP. Si les statistiques descriptives et les ANOVA ont mis de l'avant certaines tendances, les analyses de régressions multiples nous ont par ailleurs permis d'adopter un raisonnement « toutes choses égales par ailleurs » ; autrement dit, de mesurer les effets nets des types de SABP, c'est-à-dire lorsque les dépenses publiques, la variation du niveau socioéconomique des élèves et la stratification des systèmes d'éducation sont maintenues constantes.

Nos résultats suggèrent que le type de SABP caractérisé par une *accountability* de forte intensité (type 4) semble le seul qui se démarque significativement par des degrés d'efficacité plus faibles et d'inégalités des acquis de base plus importants par rapport au type de référence (le type 1, défini par une *accountability* de faible intensité).

Si les ANOVA avaient pourtant suggéré que le type 3 se démarque *a priori* par des degrés d'inégalités de résultats et des chances de résultats plus faibles que le type de référence, cette relation devient non significative lorsque les dépenses publiques en éducation, la variation du milieu socioéconomique des élèves et la stratification du système éducatif sont maintenues constantes.

En résumé, si seul le type empirique caractérisé par une *accountability* de forte intensité (type 4) se démarque dans nos analyses par des degrés d'efficacité et d'égalité des acquis de base plus faibles, il faut par ailleurs noter qu'aucune relation significative n'est constatée avec les degrés d'égalité de résultats et des chances de résultats. En somme, lorsque les dépenses publiques en éducation, la variation du niveau socioéconomique des élèves ainsi que la stratification du système scolaire sont prises en compte, ce type empirique ne semble pas faire de différence en termes d'inégalités de résultats et des chances de résultats.

Il apparaît donc, selon nos résultats, que les SABP soit n'ont pas d'effet sur différents indicateurs d'équité, soit, en ce qui concerne les systèmes d'*accountability* à forts enjeux (type 4), qu'ils entraînent des conséquences négatives, accentuant les inégalités des acquis de base dans les

systèmes d'éducation c'est-à-dire la proportion d'élèves n'ayant pas atteint, à 15 ans, un niveau de compétences de base. Les résultats de notre recherche vont donc, en partie, dans le sens des travaux menés sur les effets de l'*accountability* en éducation, plus particulièrement des recherches qui ont mis en exergue l'impact négatif des SABP à forts enjeux sur l'équité des systèmes d'éducation (Figlio & Loeb, 2011; Hanushek & Raymond, 2005; Maroy & Voisin, 2017b; Verger & Parcerisa, 2017).

En termes d'efficacité scolaire, nos résultats viennent nuancer les tendances qui ont été mises en exergue par certaines études. Plusieurs recherches sont en effet arrivées aux conclusions que les SABP à forts enjeux participeraient à une augmentation des performances moyennes des élèves, plus particulièrement en mathématiques et dans les niveaux de scolarité élevés (8^e) (Dulude & Dupriez, 2014; Mons, 2009; Mons & Dupriez, 2011). Néanmoins, comme nous l'avons discuté dans notre revue de littérature⁶⁸, ces conclusions ont été fortement nuancées par des études subséquentes, amenant ainsi à remettre en doute la capacité des SABP à forts enjeux à améliorer l'efficacité des systèmes éducatifs (voir en outre Mons, 2009 et Lee, 2010 à ce sujet). C'est aussi la conclusion à laquelle abouti notre recherche.

⁶⁸ Voir SECTION 2 : Effets des systèmes d'*accountability* basés sur la performance sur l'efficacité et l'équité scolaires, revue de la littérature

5.10 Synthèse et conclusion des résultats de recherche

En guise de synthèse, rappelons que notre premier volet empirique de recherche nous a permis de faire émerger trois composantes instrumentales des SABP d'un ensemble d'instruments clefs mis en œuvre au sein des établissements scolaires et renseignés par la base de données du PISA. Ces composantes renvoient ainsi à des configurations instrumentales qui mettent en acte des logiques de régulation différenciées et devraient *in fine* participer à l'amélioration des résultats des élèves et, plus largement, des performances du système : 1) une logique d'évaluation selon la performance; 2) une logique reposant sur la mise en place d'incitatifs directs pour les enseignants ; 3) une logique reposant sur la comparaison des performances des établissements scolaires et la mise en place de dispositifs internes visant l'amélioration de ces dernières.

Plus ou moins privilégiées au sein des systèmes éducatifs de notre échantillon, ces composantes instrumentales des SABP, ou logiques instrumentales, ont constitué la base d'une typologie empirique exploratoire des SABP. Cette typologie nous a ainsi permis de différencier quatre types de SABP : 1) une *accountability* de faible intensité et peu instrumentée; 2) une *accountability* axée principalement sur l'évaluation de la performance; 3) une *accountability* axée sur l'évaluation de la performance et la comparaison externe/amélioration interne de la performance ; 4) une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants.

Nous avons aussi analysé, dans un second volet de recherche, la relation statistique entre ces composantes instrumentales d'une part, les types empiriques de SABP d'autre part, sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs de notre échantillon. Nos résultats suggèrent premièrement que la composante se rapportant aux incitatifs articulés aux évaluations des enseignants est significativement liée à des degrés d'efficacité, d'égalité de résultats et des acquis de bases plus faibles. La composante ayant trait à la comparaison entre établissements scolaires semble elle aussi associée à une altération à la baisse du degré d'efficacité et d'égalité des acquis de base. Néanmoins, tel que nous l'avons discuté, ce résultat doit être considéré avec prudence et ne nous semble pas pouvoir être pris en considération (du fait, entre autres, des seuils de significativité à partir desquels la relation statistique se donne à voir).

Deuxièmement nos analyses suggèrent que le type de SABP qui se caractérise par une *accountability* de forte intensité prévoyant des incitatifs directs pour les enseignants (type 4), et

regroupe les systèmes éducatifs des pays de l'est de l'Europe, de l'Angleterre et de la Suède (entre autres), est associé significativement à des degrés d'efficacité et d'égalité des acquis de base plus faibles que les autres types de SABP. En toute logique, on peut en déduire que les relations entretenues par la composante instrumentale se rapportant à la mise en œuvre d'incitatifs directs pour les enseignants avec ces indicateurs d'efficacité et d'équité scolaire se retrouve dans l'analyse des effets de ce type de SABP. On pourrait aussi en déduire que la relation de la composante de comparaison entre établissements scolaires (à la limite du seuil de significativité dans les modèles intégrant les effets bruts des types de SABP), se renforce lorsqu'elle est conjuguée à celle qui se rapporte aux incitatifs à l'égard des enseignants. Toutefois cette dernière déduction mériterait d'être mise à l'épreuve.

En résumé nous pourrions donc statuer, à la lumière de ces résultats, que l'hypothèse d'un effet différencié des SABP (composantes instrumentales et types empiriques) est partiellement validée par nos analyses, puisqu'une composante instrumentale (incitatifs) et un type d'*accountability* (type 4) se distinguent par des relations statistiquement significatives, négatives au demeurant, avec les degrés d'efficacité et d'égalité des acquis de base⁶⁹ (mais aussi avec le degré d'égalité de résultats dans le cas de la composante des incitatifs). Cette conclusion est toutefois à nuancer.

Il faut premièrement prendre en compte les tailles d'effets des modèles qui examinent l'impact (effet net) des composantes instrumentales et des types empiriques de SABP sur l'efficacité et les différents indicateurs d'équité scolaires. Dans le cas de la composante des incitatifs, les coefficients de détermination des modèles qui interrogent l'effet net de cette dernière oscillent entre environ 24% et 38% de la part de variation des degrés d'efficacité, d'inégalités de résultats et des acquis de base dans les systèmes éducatifs d'Europe et du Canada. Nous avons cependant vu que l'introduction des variables de contrôle engage une augmentation sensiblement importante des coefficients de détermination des modèles (de 20% à 30%).

Dans le cas du type 4 de SABP, le constat est sensiblement le même. Les tailles d'effets des modèles sont relativement faibles lorsque les variables de contrôle en sont exclues (moins de 10%). L'introduction de ces dernières engage en effet une augmentation relativement importante des coefficients de détermination des modèles (entre 18% et 17%).

⁶⁹ Si la composante de comparaison entre établissements semble elle aussi associée significativement à ces dimensions, elle l'est toutefois à des seuils de significativité supérieurs à 0,05, et avec une taille d'effet des modèles très faible, ce qui nous a amenée à émettre de fortes réserves quant à ce résultat.

Deuxièmement, il faut relever le fait que ni les composantes instrumentales ni les types empiriques de SABP ne semblent altérer les degrés d'inégalité de résultats ou des chances de résultats. Les effets des composantes instrumentales et des types empiriques de SABP ne se donnent donc pas à voir dans les mêmes termes en fonction des conceptions de l'équité, et donc des inégalités qui sont examinées.

Troisièmement il faut aussi rappeler que les analyses préliminaires ont mis en exergue le fait que les relations entre composantes instrumentales, types empiriques de SABP, efficacité et équité scolaires sont soit très faibles, soit inexistantes, en fonction des composantes et des types de SABP considérés et des dimensions étudiées (efficacité et équité). Nos résultats aboutissent donc, au total, à des conclusions plus que mitigées sur les effets des SABP sur l'efficacité et l'équité scolaires.

Par contraste, il faut mettre en exergue l'influence relativement stable de la variation du niveau socioéconomique des élèves et de la stratification des systèmes éducatifs ainsi que des dépenses publiques en éducation sur les indicateurs d'efficacité et d'équité mobilisés.

La variation du niveau socioéconomique des élèves est en effet systématiquement associée à des degrés d'efficacité plus faibles, et à des degrés d'inégalité de résultats et des acquis de base plus élevés. Ces résultats vont donc dans le sens du lien, nombre de fois discuté dans la littérature, entre niveau socioéconomique des élèves et performances scolaires, que les SABP, selon nos analyses, ne permettent pas d'altérer.

La stratification des systèmes scolaires est, quant à elle, associée systématiquement à des degrés d'inégalités des chances de résultats plus élevés, mais aussi à des degrés d'inégalités de résultats plus importants, constat généralement mis en exergue par la littérature. Nous avons par ailleurs constaté que ce facteur était aussi associé à des degrés d'efficacité scolaire élevés. Une déduction possible, mais qui mérite toutefois d'être mise à l'épreuve, serait que la stratification des systèmes scolaires tire les meilleurs élèves vers le haut, leurs performances se donnant à voir dans une augmentation des performances moyennes au sein des systèmes éducatifs, mais participe par ailleurs à la discrimination des élèves les plus faibles et les moins favorisés au niveau socioéconomique.

Finalement, l'effort budgétaire que consentent les pays ou les juridictions dans l'éducation est, selon nos analyses, aussi associé à des degrés d'inégalités des chances de résultats plus faibles,

bien qu'à des seuils de significativité dépassant le seuil de 0,05. En d'autres termes, il participerait à la réduction de la force du lien entre niveau socioéconomique des élèves et leurs performances scolaires, constat qui émerge par ailleurs de la littérature.

CHAPITRE 6 : DISCUSSION GÉNÉRALE

Si le chapitre précédent a permis d'une part de caractériser les SABP dans les systèmes éducatifs d'Europe et du Canada, d'autre part d'analyser leurs effets sur les degrés d'efficacité et d'équité scolaires de ces systèmes d'éducation, ce dernier chapitre propose une discussion des résultats de notre recherche doctorale, ainsi qu'une réflexion plus large sur les SABP et les tendances actuelles en termes de politiques éducatives. Ainsi nous discutons premièrement chacun des volets empiriques de notre recherche. Nous résumons les principaux résultats qui en émergent pour les mettre en perspective et en dialogue avec la littérature pertinente. Ceci nous amène par la suite à questionner la rhétorique qui entoure la mise en œuvre des SABP à la lumière de nos résultats de recherche. Finalement, nous dépassons le cadre de nos résultats de recherche pour mettre en débat certains éléments qui nous semblent clefs lorsque l'on s'intéresse à la mise en œuvre des SABP dans les systèmes éducatifs, et plus largement à l'avènement de la régulation des systèmes éducatifs par leurs résultats. Ainsi, nous nous interrogeons sur les effets normatifs des SABP en nous appuyant sur la littérature d'orientation critique. Finalement, après avoir proposé une discussion des principes de justice au cœur des SABP, nous élargissons la discussion vers une autre voie possible pour penser les politiques de régulation par les résultats et les SABP dans les systèmes d'éducation.

6.1. Discussion du premier volet de recherche : caractériser les systèmes d'*accountability* basés sur la performance dans les systèmes éducatifs d'Europe et du Canada

L'analyse comparative des SABP dans les systèmes éducatifs d'Europe et du Canada visait à caractériser et analyser ces derniers à partir des instruments centraux qui en constituent le socle.

Si le développement des SABP s'inscrit dans un contexte de transformation des modes de régulation mis en place par l'État pour orienter et coordonner l'action éducative, notre recherche s'est inscrite dans une conception des SABP comme forme de régulation formelle qui s'appuie sur un ensemble d'instruments visant à piloter les systèmes éducatifs par leurs résultats.

Notre premier volet de recherche visait donc à caractériser les SABP dans les systèmes éducatifs de notre échantillon sur base des instruments mis en œuvre au sein des établissements scolaires pour cadrer et orienter le comportement des acteurs. Si la littérature sur les instruments d'action

publique met en exergue le fait que les instruments sont porteurs d'une certaine conception de la régulation (Lascoumes & Legalès, 2004; Lascoumes & Simard, 2011), nous avons fait le postulat que des instruments différenciés sous-tendent finalement des logiques de régulation distinctes et permettent de distinguer plusieurs types de SABP.

Les variations dans la gamme d'instruments privilégiés dans les différents systèmes scolaires de notre échantillon nous ont ainsi permis de faire émerger des logiques de régulation au cœur des SABP qui, par ailleurs, ont pour point commun de chercher à mettre en place - à des degrés et selon des modalités diverses - une régulation en fonction des performances des écoles ou du système. Autrement dit, grâce à notre analyse en composante principale, nous avons empiriquement dégagé, à partir d'une série de variables renseignées par la base de données PISA 2012, des « composantes instrumentales » des SABP, en d'autres termes des logiques instrumentales de régulation constitutives des systèmes d'*accountability*. Nous ne partions pas *a priori* de ce que devait être une logique de régulation par les résultats, mais bien des instruments mis en œuvre au sein des établissements scolaires dans les systèmes éducatifs de notre échantillon. Nous avons aussi suggéré que ces logiques instrumentales de régulation pouvaient reposer sur des théories de l'action sensiblement différentes, bien que non mutuellement exclusives.

Les analyses de *cluster* nous ont en effet permis d'observer comment ces composantes instrumentales sont articulées et combinées dans les faits (selon des modalités différenciées) au sein de types empiriques de SABP. Sur base des composantes instrumentales extraites nous avons ainsi élaboré une typologie exploratoire des SABP qui permettent de différencier les systèmes éducatifs d'Europe et du Canada.

6.1.1 Des instruments aux logiques instrumentales des systèmes d'*accountability* basés sur la performance

Nos analyses ont ainsi premièrement fait émerger trois composantes principales interprétées comme trois logiques instrumentales de régulation : 1) une logique d'évaluation et de suivi de la performance; 2) une logique reposant sur l'articulation de conséquences aux évaluations des enseignants; 3) une logique s'appuyant sur la comparaison des performances entre établissements et des dispositifs visant à soutenir le changement de pratiques au sein des établissements scolaires.

- L'évaluation et le suivi de la performance des acteurs ; la première composante instrumentale extraite de nos analyses renvoie à une dimension centrale des SABP : l'évaluation et le suivi de la performance des acteurs. Elle repose sur un ensemble d'instruments et de mécanismes permettant d'évaluer et de suivre les résultats des acteurs (élèves, enseignants, établissements scolaires), dans une logique qui fait du suivi de la performance la clef du changement organisationnel.

Il apparaît que cette logique est fortement développée dans les régions espagnoles (Madrid, Andalousie, Catalogne) et dans certaines provinces canadiennes (Terre-Neuve et Labrador, Alberta) ou encore en Écosse, mais peu dans les *Länders* allemands inclus à notre échantillon, en Grèce, en Italie, en France ou encore dans les communautés francophones et germanophones de Belgique.

- Le lien évaluation des enseignants –conséquences ; la seconde composante instrumentale repose sur l'articulation de conséquences de différentes natures aux évaluations du corps enseignant. Ces conséquences matérielles et sociales constituent ainsi le cœur d'une logique qui fait des incitatifs (pour les acteurs enseignants) les médiations privilégiées du changement.

Cette dernière serait davantage privilégiée dans les pays de l'est de l'Europe tels que la Slovaquie, la Hongrie, la République tchèque et la Pologne, mais aussi en Angleterre et en Suède. Elle s'avère, à l'inverse, relativement marginale dans les provinces canadiennes de l'Ontario, de la Nouvelle-Écosse, de Terre-Neuve et Labrador, de la Colombie-Britannique, mais aussi dans les régions espagnoles d'Aragon, de la Rioja, en Galice, ou encore en Irlande, en Belgique francophone et germanophone.

- Comparaison externe des performances et dispositifs internes d'incitation au changement ; la dernière composante réfère, quant à elle, à une logique instrumentale qui s'appuie sur l'action supposée de la comparaison externe des performances de l'établissement avec d'autres performances de référence (d'autres établissements ou une entité administrative englobante), comparaison qui peut induire de la concurrence entre établissements scolaires. De plus, cette comparaison externe est couplée avec la mise en place de dispositifs formels d'intervention au sein des établissements scolaires, visant à

soutenir le changement organisationnel et à inciter les acteurs à adopter de nouvelles pratiques plus efficaces.

D'après nos résultats, c'est dans les systèmes éducatifs de certaines provinces canadiennes telles que l'Ontario, la Saskatchewan et la Nouvelle-Écosse, et dans les systèmes éducatifs européens de l'Angleterre, du Pays de Galles, de l'Irlande du Nord, de l'Écosse que cette logique est davantage développée. À l'opposé, elle est de plus faible intensité dans les régions d'Espagne de la Galice, de Castille et Léon, de la Murcia ou encore dans les *Länders* allemands de la Basse-Saxe ou de l'Hesse.

6.1.2 Les logiques instrumentales comme base d'une typologie exploratoire des systèmes d'*accountability* basés sur la performance

Ces logiques instrumentales mobilisées comme base de l'analyse de classification nous ont par ailleurs permis de faire émerger quatre types empiriques de SABP au sein desquels sont regroupés les systèmes éducatifs d'Europe et du Canada.

1) Le premier type de SABP se définit par une *accountability* de faible intensité et peu instrumentée. La gamme d'instruments mise en œuvre au sein des établissements scolaires y est sensiblement plus restreinte que dans les autres types empiriques. Il permet de caractériser une large part des systèmes éducatifs de notre échantillon (20 systèmes éducatifs, dont les *Lands* allemands, la France, la Suisse, l'Autriche, la Belgique, Luxembourg, du Danemark, de la Norvège, de l'Irlande, les provinces canadiennes du Québec et du Manitoba, etc.).

2) Le deuxième type de SABP se définit quant à lui par une *accountability* qui repose sur une logique et des instruments axés principalement sur l'évaluation de la performance. Il permet de caractériser les régions d'Espagne, la Finlande, le Portugal et l'Islande (soit 17 systèmes éducatifs au total). Nos analyses suggèrent que la régulation du système opère dans ce type de SABP principalement au moyen d'un ensemble d'instruments visant à contrôler et à suivre les performances des établissements scolaires et des élèves. Toutefois, à l'inverse du type suivant (type 3), la logique de comparaison entre établissements scolaires y est très peu développée, de même que celle reposant sur la mise en place d'incitatifs directs pour les enseignants.

3) Le troisième type de SABP, qui caractérise la majorité des provinces canadiennes et l'Écosse (soit 8 systèmes d'éducation au total), articule quant à lui logique instrumentale d'évaluation

selon la performance et logique de comparaison externe/amélioration interne des performances des établissements scolaires. Il repose sur une gamme d'instruments étendue permettant d'une part d'opérer un suivi serré des performances des établissements, d'autre part de stimuler la concurrence entre établissements scolaires (par la comparaison de leurs résultats) et de rendre ces derniers plus efficaces par la mise en place de dispositifs d'intervention en leur sein. Néanmoins, à l'instar des types de SABP précédents et à l'inverse du type suivant (le type 4) les incitatifs directs pour les enseignants y sont peu valorisés.

4) Finalement, le dernier type de SABP articule plusieurs logiques instrumentales et regroupe les systèmes éducatifs des pays de l'est de l'Europe (la Bulgarie, l'Estonie, la Roumanie, la Pologne, etc.), avec ceux de l'Angleterre, de la Suède et des Pays-Bas (14 systèmes d'éducation au total). Cette *accountability* de forte intensité prévoit une large gamme d'instruments et accorde une place de choix aux incitatifs directs pour les enseignants, alors que ces dispositifs sont marginaux dans les autres types empiriques. Ainsi c'est l'importance accordée à cette dimension qui le distingue clairement des autres types de SABP. Par ailleurs la logique de comparaison entre établissements et d'amélioration de leurs performances par le recours à des dispositifs internes (intervention d'experts externes et activités de développement professionnel offertes aux enseignants) y est sensiblement moins développée qu'au sein du type précédent (type 3).

Au total, nos analyses nous ont donc permis de mettre en exergue, d'une part, des logiques instrumentales de régulation différenciées, et d'autre part, de quelle manière ces dernières s'articulent au sein de types empiriques de SABP. Ces logiques et les instruments qui les mettent en oeuvre sont tantôt significatifs, pour reprendre les propos de Barrère (2011, p. 127), de « l'alliance prescrite de l'évaluation et de l'action locale », et des efforts entrepris par les autorités institutionnelles pour mesurer les performances des acteurs, responsabiliser ces derniers ou les rendre imputables quant aux résultats des élèves (Harris & Herrington, 2006). Elles s'inscrivent dans une volonté de soumettre les établissements scolaires à la pression de la compétition stimulée par la mise en comparaison de ces derniers sur base de leurs performances.

Non mutuellement exclusives, articulées selon les cas au sein des types empiriques de SABP, ces logiques de régulation s'appuient ainsi sur différents instruments visant l'évaluation et le suivi des performances des acteurs, sur des mécanismes de comparaison des performances des établissements scolaires et des dispositifs d'intervention en leur sein, sur des incitatifs formalisés

pour les acteurs enseignants qui constituent dès lors les clefs du changement organisationnel. Elles s'appuient ainsi sur un ensemble d'instruments qui, tel que la littérature l'a montré, s'avèrent clefs pour analyser l'*accountability* en éducation et comprendre les modalités d'opérationnalisation de la régulation des systèmes éducatifs par leurs résultats.

6.1.3 Logiques et types empiriques de systèmes d'*accountability* basés sur la performance : la question des enjeux pour les acteurs

Si nous avons souhaité dépasser la distinction empirique classique *accountability* à forts et à faibles enjeux (Carnoy & Loeb, 2002), *accountability* douce versus *accountability* dure (Dupriez & Mons, 2011), force est de constater que la dimension des enjeux s'avère aussi clef dans l'analyse des SABP dans les systèmes éducatifs d'Europe et du Canada.

Cette dimension peut ainsi premièrement être appréhendée au travers de la composante qui se rapporte aux incitatifs articulés aux évaluations des enseignants. Cette logique instrumentale permet en effet de saisir les conséquences formelles auxquelles font face une catégorie d'acteurs clefs, les enseignants, et d'observer les systèmes éducatifs qui prévoient donc d'arrimer des enjeux à leur évaluation.

Cette composante instrumentale s'est avérée en outre clef dans la distinction des types empiriques de SABP qui émergent de nos analyses. D'une part, elle discrimine leur formation (cette composante instrumentale a en effet le poids le plus important dans la formation des *clusters*) et, d'autre part, permet clairement de distinguer un type de SABP des autres types émergents. Ainsi, le type de SABP caractérisé par une *accountability* de forte intensité (type 4), est le seul qui privilégie la mise en place d'incitatifs directs (et donc d'enjeux forts) pour les enseignants.

Toutefois, la question des enjeux pour les acteurs ne peut être limitée aux incitatifs visant les enseignants. On peut en effet faire l'hypothèse que plus la gamme d'instruments qui constituent le socle des SABP est étendue, plus les acteurs du champ font face à des enjeux importants et sont soumis à des pressions liées aux logiques instrumentales documentées par notre étude : pression à la performance, liée aux mécanismes de compétition entre écoles, introduite par les conséquences matérielles et sociales visant les enseignants.

Aussi, tel que la littérature l'a montré, des conséquences symboliques et réelles pour les acteurs sont en effet liées à différents modes de régulation mis en place au sein des établissements scolaires. À titre d'exemple, la publication des résultats des élèves et des écoles induit une forte pression sur les établissements scolaires et présente des enjeux forts pour ces derniers, particulièrement dans un contexte de marché scolaire. Les études ont ainsi mis en exergue le phénomène de *public shame* lié à la publication des résultats des écoles, en d'autres termes la stigmatisation des établissements scolaires les moins performants et des professionnels qui y œuvrent (Carnoy & Loeb, 2002; Wolf & Jansens, 2007). Les performances des écoles constituent ainsi l'un des critères à partir desquels la qualité des établissements est jugée par le public, par les autorités administratives, et donc à partir desquelles l'image externe des établissements scolaires est construite. Au-delà des conséquences symboliques, ce phénomène entraîne par ailleurs des conséquences réelles ayant trait à la difficulté pour ces écoles d'attirer et de retenir un personnel éducatif qualifié. Ceci nous invite donc à mettre en perspective la question des enjeux pour les acteurs de manière plus large.

Si nous prenons en compte ces éléments, il apparaît dès lors que nous pourrions centrer l'analyse (et la différenciation) des types empiriques de SABP qui émergent de nos analyses à la lumière de cette dimension des enjeux. Ainsi, les types de SABP qui s'appuient sur une gamme d'instruments étendue et qui articulent plusieurs composantes instrumentales, conjuguant ainsi plusieurs modes de régulation et mécanismes de reddition de comptes, revêtent finalement des enjeux sensiblement plus importants que les types de SABP où les gammes d'instruments sont plus restreintes. Nous pourrions dès lors situer les types empiriques construits sur un continuum allant des SABP à plus faibles enjeux vers les SABP à forts enjeux.

Toutefois, au-delà de la force des enjeux (enjeux forts versus enjeux faibles), il nous paraît essentiel de prendre en compte la nature sensiblement différente de ces derniers qui varie en fonction des types de SABP, des instruments et des logiques de régulation qui y sont privilégiés. Ils peuvent ainsi être plus ou moins formalisés, s'adresser aux acteurs à titre individuel ou collectif, prendre la forme de conséquences indirectes, de sanctions symboliques liées à différents mécanismes tels que la compétition entre établissements scolaires. Cette distinction ayant trait à la nature des enjeux et des conséquences pour les acteurs (écoles, enseignants, administrateurs scolaires) que la catégorisation enjeux forts - enjeux faibles ne permet pas de saisir, nous semble en effet centrale pour différencier les SABP dans divers systèmes éducatifs.

6.2 Discussion du second volet de recherche : effets des systèmes d'*accountability* basés sur la performance sur l'efficacité et l'équité scolaires, un constat mitigé

Au total, si notre étude a permis de mettre en exergue différentes logiques instrumentales et différents types empiriques de SABP, il s'agissait d'observer dans quelle mesure ces derniers permettent, le cas échéant, de rendre les systèmes éducatifs plus efficaces et plus équitables.

Les analyses menées dans un second volet de recherche empirique ont ainsi examiné les relations entre les logiques instrumentales, les types empiriques de SABP et les degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada. Puisque les considérations d'efficacité et d'équité scolaires entourent le développement des SABP dans la majorité des systèmes éducatifs, nous avons souhaité analyser ces derniers à la lumière des finalités qu'ils poursuivent officiellement.

Nous avons donc fait l'hypothèse, basée sur la littérature et au cœur de ce deuxième volet empirique de recherche, d'un effet différencié des SABP sur les degrés d'efficacité et d'équité scolaires en fonction des logiques instrumentales privilégiées et des types empiriques qu'ils revêtent.

6.2.1 Systèmes d'*accountability* basés sur la performance, efficacité et équité : des effets qui demeurent faibles et instables

D'une manière générale, notre recherche nous amène à dresser un constat plus que mitigé quant aux potentialités des SABP. Les relations entre SABP (logiques instrumentales et types empiriques), efficacité et équité scolaires sont en effet soit d'une intensité très faible, soit non significatives dans la plupart des cas. Seuls les incitatifs pour les enseignants et le type empirique (type 4) privilégiant la mise en œuvre de ce dispositifs (articulés par ailleurs à d'autres modes de régulation tel que le suivi des résultats des élèves au niveau des établissements scolaires, l'évaluation des établissements scolaires, etc.) se distinguent par des relations significatives avec les degrés d'efficacité et d'équité scolaires. Dans les deux cas, l'effet constaté est négatif, impliquant que ces derniers (cette composante instrumentale et ce type de SABP) sont facteurs d'inefficacité et d'iniquité, avec cependant des variations en fonction des conceptions de l'équité, donc des figures de l'égalité qui sont étudiées.

En effet, d'après nos analyses, les incitatifs pour les enseignants sont liés à des degrés d'efficacité, d'égalité de résultats et des acquis de base plus faibles dans les systèmes éducatifs de notre échantillon. Le constat est sensiblement similaire s'agissant du type de SABP (type 4) se rapprochant davantage de ce que littérature nomme SABP à forts enjeux : bien que ce type ne semble pas avoir d'effet sur l'égalité de résultats entre élèves (à la différence de la composante instrumentale d'incitatifs directs pour les enseignants), il est lui aussi lié à des degrés d'efficacité et d'égalité des acquis de base plus faibles.

Nos résultats, qui doivent être considérés à la lumière des limites de notre recherche (qui seront discutées plus avant), s'inscrivent ainsi dans la lignée des études qui remettent en question les effets positifs des SABP à forts enjeux sur l'efficacité scolaire (Mons, 2009; Lee, 2010). Comme nous l'avons vu, certaines recherches aboutissent en effet aux conclusions d'un effet positif de systèmes d'*accountability* à forts enjeux sur les performances des élèves, principalement dans les niveaux de scolarité supérieurs et en mathématiques (Hanushek & Raymond, 2005; Jacob, 2005; Harris & Herrington, 2006; Woessman, 2007 ; Lee, 2008; Figlio & Loeb, 2011). Toutefois, les divergences de conclusions entre ces études et d'autres études subséquentes au même titre que la notre (Lee, 2008; 2010 ; Chiang, 2009; Mons, 2009) invite à émettre de sérieux doute quant à la capacité des SABP, et plus particulièrement des SABP à forts-enjeux, à améliorer l'efficacité scolaire.

En termes d'équité, les tendances qui émergent de la littérature pointent dans le sens de nos résultats, (Hanushek & Raymond, 2005; Harris & Herrington, 2006; Lee, 2008; Figlio & Loeb, 2011). En effet nombreuses sont les études s'intéressant aux effets des SABP sur l'équité scolaire qui ont mis en exergue que ces derniers participeraient de l'augmentation de l'écart de performances entre groupes d'élèves (Dee, 2002 cité dans Carnoy & Loeb, 2002). Dans le meilleur des cas, si les systèmes d'*accountability* à forts enjeux ne creusent pas l'écart entre groupes d'élèves, leur capacité à améliorer l'équité est fortement remise en question (Amrein & Berliner, 2002 : Lee & Wong, 2004 ; Lee, 2010).

Les résultats de notre étude suggèrent que la composante se rapportant à la mise en place d'incitatifs directs pour les enseignants participerait de l'augmentation de l'écart de résultats entre les élèves les plus faibles et les élèves les plus performants d'une part, mais aussi de la proportion d'élèves n'ayant pas acquis, à 15 ans, un niveau de compétence de base (le niveau 2

du PISA). Par ailleurs la proportion d'élèves n'ayant pas acquis un niveau de compétence de base est aussi plus importante dans les systèmes éducatifs ayant mis en place des SABP à forts enjeux, c'est à dire des SABP caractérisés par une *accountability* de forte intensité reposant, entre autres, sur des incitatifs directs pour les enseignants (type 4).

La littérature permet de proposer des pistes explicatives aux tendances qui émergent de notre thèse doctorale. Il ressort en effet de plusieurs études que les systèmes d'*accountability* à forts enjeux, marqués en outre par une forte compétition entre écoles, provoqueraient des phénomènes qui pénalisent les publics les plus fragilisés : focalisation des ressources éducatives sur les élèves les plus à même d'améliorer leurs performances au détriment des plus faibles, stigmatisation des élèves et des écoles sous performants, amplification de la ségrégation académique et sociale, mise en place, par ces dernières, de solutions à court terme qui finalement s'avèrent contre-productives, difficultés pour les établissements problématiques d'attirer et de retenir un personnel qualifié (Wolf & Janssens, 2007; Finnigan & Gross, 2007; Mintrop & Sunderman, 2009 ; Figlio & Loeb, 2011). On peut dès lors supposer que des phénomènes de ce type sont potentiellement à l'œuvre dans les systèmes éducatifs de notre échantillon ayant choisi la voie de l'*accountability* à forts enjeux. C'est, en tout cas, ce que suggèrent les études de Wolf et Janssens (2007), Rustique-Forrester (2005) et de Rosenthal (2004) dans le contexte anglais. Des recherches subséquentes seraient donc nécessaires pour approfondir les tendances mises en exergue par notre recherche afin mieux comprendre les processus par lesquels les SABP à forts-enjeux altèrent l'équité scolaire des éducatifs ayant choisi cette voie de l'*accountability*.

Précisons toutefois que, dans le cadre de notre recherche, nous ne pouvons pas exclure la possibilité de la présence d'une relation de causalité inversée au sein des systèmes éducatifs considérés. La relation s'inverse alors et les effets précèdent les causes. En d'autres termes, les systèmes éducatifs les moins performants et les plus inéquitables de notre échantillon pourraient être sujets à des pressions plus importantes pour mettre en place des SABP à forts enjeux et des incitatifs directs pour les enseignants. Si c'est une possibilité que nous ne pouvons exclure, elle n'invalide pas pour autant nos résultats, mais demeure une question qui mérite d'être posée.

Dans le contexte des États-Unis, Carnoy et Loeb (2002) ont en effet suggéré que les États enregistrant les taux de performances les plus faibles, ainsi qu'une forte proportion d'élèves issus des minorités ethniques, sont aussi ceux qui ont été plus à même de mettre en place des SABP

prévoyant des enjeux forts pour les acteurs. Cette question devrait donc être investiguée plus avant dans le contexte des systèmes éducatifs d'Europe et du Canada, plus particulièrement au sein de ceux qui privilégient la mise en œuvre d'incitatifs directs pour les enseignants, et qui ont adopté des SABP à forts enjeux, comme les pays de l'est de l'Europe jusqu'à présent peu étudiés.

6.2.2 Relation entre systèmes d'*accountability* basés sur la performance et réduction des inégalités scolaires d'origine socioéconomique : des effets inexistantes

Par ailleurs, il faut aussi observer le fait que les effets des SABP mis en avant par nos analyses sont faibles et varient en fonction des conceptions de l'équité étudiées. L'un des résultats saillants de notre recherche est ainsi qu'aucune relation n'est constatée entre les SABP et le degré d'inégalités des chances de résultats. Ce constat, qui va dans le sens de l'étude de l'OCDE menée à partir des données du PISA 2006 (2007 citée dans Mons, 2009), implique que les SABP ne parviennent pas à atténuer la force du lien entre niveau socioéconomique des élèves et performances scolaires. Si cet enjeu clef a fait, et fait encore, l'objet de nombreux développements dans la littérature, les potentialités des SABP sont ici limitées, voire inexistantes.

Le constat, fort discuté dans la littérature, d'une forte relation entre niveau socioéconomique des élèves et équité scolaire (Coleman, 1966; Jencks, 1972; Mons, 2009; OCDE, 2000; 2003; 2009; 2012, etc.) reste valide au-delà des conceptions de l'équité examinées dans notre recherche. Les mécanismes par lesquels les ressources socioculturelles et économiques impactent la réussite et le parcours scolaires des élèves, et par lesquels les avantages et les désavantages sociaux se convertissent progressivement en désavantages et avantages scolaires (Bourdieu, 1966), sont aujourd'hui bien connus, notamment grâce aux travaux de la sociologie de l'éducation. Si certains systèmes éducatifs arrivent cependant, mieux que d'autres, à atténuer l'impact de l'origine socioéconomique sur les résultats scolaires et que, pour reprendre les propos de Dubet (Dubet, 2010), l'école ne transforme pas les inégalités sociales en inégalités scolaires de manière constante, les SABP ne semblent pas ici jouer un rôle déterminant.

Si l'école peut faire la différence dans la réduction des inégalités scolaires d'origine sociale, notre étude ne permet que partiellement de discuter la proposition, fort débattue dans la littérature, selon laquelle les « bonnes pratiques » et l'*accountability* permettent effectivement

d'atténuer le lien entre niveau socioéconomique des élèves et performances scolaires. Si notre étude nous permet de soutenir que l'*accountability* à elle seule ne semble pas contribuer à la réduction de la détermination socioéconomique des performances scolaires, il s'agirait d'observer dans quelle mesure l'*accountability* et ses instruments influencent les pratiques mises en œuvre au sein des établissements (pratiques de gestion et pratiques éducatives), qui à leur tour affecteraient potentiellement l'efficacité et l'équité scolaires, pour pouvoir nous positionner plus largement au sein de ce débat.

6.3 Stratification des systèmes éducatifs et dépenses publiques en éducation : freins ou vecteurs d'équité scolaire et d'efficacité scolaire

D'après les résultats de notre recherche, qui confirment en partie ceux de la littérature, d'autres facteurs s'avèrent clefs dans la réduction des inégalités scolaires : les dépenses publiques investies dans l'éducation ainsi que la stratification horizontale et verticale du système. Si les dépenses publiques en éducation sont liées à une réduction de la détermination socioéconomique des élèves, constat déjà réalisé par Mons (2009), la stratification du système scolaire entraîne quant à elle une dégradation de l'équité, plus particulièrement de l'égalité de résultats et des chances de résultats.

Sur la question des effets de la stratification sur l'équité scolaire, les résultats de la littérature sont sans appel (Duru-Bellat, Jarousse, & Mingat, 1993; Duru-Bellat & Mingat, 1988; Duru-Bellat et al., 2004; Mons, 2004, 2007; Vandenberghe, 2001). Si les choix d'orientation, les aspirations des élèves et de leurs familles reposent sur des critères « objectifs » tels que les compétences académiques (ces dernières étant généralement fortement liées au niveau socioéconomique des élèves), ces choix sont aussi fortement influencés par des facteurs socioculturels dont résultent des ambitions et les aspirations scolaires. Cette proposition reste vraie que l'on se place dans le cadre d'une analyse en termes de domination sociale (Bourdieu & Passeron, 1970) ou faisant la part belle à l'autonomie (relative) des acteurs et à leur choix selon Boudon (Boudon, 1973). Ces mécanismes jouent aussi dans les questions des « choix » en termes de redoublement ou d'établissements scolaires (Van Zanten, 2009a, 2009b).

Toutefois, nos résultats suggèrent que la stratification du système serait liée à des degrés d'efficacité scolaire plus élevés dans les systèmes éducatifs de notre échantillon. Nous pourrions en déduire qu'en favorisant le triage des élèves au sein du système éducatif, la stratification du

système aurait dès lors comme effet de tirer les meilleurs élèves vers le haut, leurs performances affectant la moyenne des scores, et d'avoir l'effet inverse sur les élèves les plus faibles. Cependant, le lien entre stratification du système et efficacité scolaire, plus particulièrement la présence de filières différenciées, qui scellent les parcours scolaires vers des carrières scolaires plus ou moins prestigieuses (exigeantes académiquement, valorisées scolairement et socialement), a été fortement remis en question (Mons, 2008). Il s'agirait donc d'analyser plus avant cette relation dans le contexte des SABP et de questionner la possibilité du renforcement du triage des élèves au sein des différentes filières dans un contexte d'enjeux forts pour les acteurs liés aux performances des élèves.

Enfin, ces éléments renvoient à l'importance de facteurs structurels qui agissent et déterminent en partie la capacité des systèmes éducatifs à réduire le lien entre inégalités sociales et inégalités scolaires, mais aussi à mener tous les élèves vers certains niveaux d'acquisition. Ce constat ne permet pas de trancher le débat autour du poids des structures versus le comportement des acteurs qui traverse le champ et les travaux de la sociologie de l'éducation. Si nous devons résumer notre posture sur cette question nous dirions que nous pensons que la structure des systèmes (sociale, organisationnelle) joue un rôle déterminant dans l'orientation du comportement des acteurs qui bénéficient toutefois d'une marge de manœuvre « 'pour jouer' avec le système et mobiliser leurs ressources sociales afin d'en réinterpréter le sens » (Schnapper, 1999, p.13).

Or, la gouvernance par les résultats et les politiques d'*accountability* n'ont pas pour ambition d'agir sur ces facteurs structurels qui semblent pourtant fortement déterminants des degrés d'efficacité et d'équité scolaires. Il s'agit davantage de mettre en place un ensemble de dispositifs pour réguler le comportement des acteurs au sein même des structures scolaires. Dès lors, si les politiques d'*accountability* visent à contraindre encore davantage la marge de manœuvre des acteurs dans une logique d'efficacité, il faudrait aussi observer dans quelle mesure les instruments d'*accountability* orientent effectivement les comportements des acteurs vers l'adoption de pratiques supposées plus efficaces, et comment les acteurs utilisent cette marge de manœuvre pour œuvrer vers plus d'efficacité et d'équité scolaires⁷⁰.

⁷⁰ Plusieurs travaux ont été réalisés et sont en cours sur cette question : voir par exemple les travaux de Coburn (2004) dans le contexte des États-Unis, de Maroy, Mathou et Vaillancourt (Maroy, Vaillancourt, & Mathou, 2017) dans le contexte Québécois.

6.4 La rhétorique politique à l'épreuve de résultats factuels

Au total, si les politiques d'*accountability* s'appuient sur une rhétorique commune, celle de la qualité, de l'efficacité et de l'équité, les résultats de notre recherche nous amènent à mettre en doute la capacité supposée des SABP à produire les résultats escomptés. Dans le contexte des systèmes éducatifs d'Europe et du Canada en tout cas, les SABP ne semblent pas en mesure de tenir leurs promesses d'efficacité et d'équité. Les résultats de notre recherche doctorale nous conduisent donc à répondre par la négative à notre question centrale : les SABP ne permettent pas de rendre les systèmes éducatifs plus efficaces et plus équitables. Notre étude met ainsi davantage en exergue les limites de ces dispositifs qui, au-delà de leurs promesses, présentent des zones d'ombre et soulèvent de vifs questionnements.

Si la popularité des SABP auprès des décideurs politiques ne semble pouvoir être remise en question, tout se passe comme si le bien-fondé de ces politiques « comme voie d'amélioration du système » (Dutercq & Maroy, 2017, p. 25) n'était pas questionné, le débat se centrant plutôt sur les modalités de mise en œuvre des SABP et les formes qu'ils revêtent.

Si l'on peut opposer que les SABP prennent des formes variées qui dépendent des contextes politiques et institutionnels au sein desquels ils sont mis en œuvre, notre recherche a justement tâché de prendre en compte la variété des SABP et des instruments sur lesquels ils reposent dans les contextes des systèmes éducatifs d'Europe et du Canada. Il s'agissait de dépasser le cadre des États-Unis où nombre de recherches ont été menées, d'observer les effets potentiellement différenciés des SABP en fonction des formes qu'ils revêtent dans ces contextes.

Or, notre recherche suggère que la capacité des SABP et des mécanismes sur lesquels ils reposent, d'améliorer l'efficacité et l'équité scolaires ne va pas de soi. Dans le même sens, les constats qui émergent de la littérature ne sont pas si univoques. Tout au plus, malgré des divergences de résultats, les effets constatés s'avèrent relativement faibles. De fait, les résultats de notre recherche, à la suite d'autres, mettent à l'épreuve la rhétorique qui entoure le développement de ces politiques à la lumière de résultats factuels.

On peut dès lors se demander si la portée des SABP ne se situe pas finalement au-delà des résultats qu'ils sont censés produire, en d'autres termes des *outputs* attendus. Si l'on accepte le postulat de selon lequel les politiques publiques « constituent des espaces au sein desquels les acteurs concernés vont construire et exprimer 'un rapport au monde' qui renvoie à la manière

dont ils perçoivent le réel, leur place dans le monde et ce que le monde devrait être » (Muller, 2000, p. 196), cela nous amène à interroger les politiques d'*accountability* dans des termes qui dépassent la question, bien que cruciale, de leurs effets en termes d'équité et d'efficacité scolaires.

Puisque les politiques publiques construisent des interprétations du réel et définissent des modèles normatifs d'action, les effets normatifs non moins cruciaux des politiques de régulation par les résultats mériteraient d'être examinés. Autrement dit, par-delà l'interrogation de cette thèse sur les effets « officiellement » visés de ces politiques que les études à grandes échelles permettent de saisir, il faudrait s'interroger sur d'autres effets symboliques et sociaux que ne peuvent appréhender ces enquêtes. L'étude de ces effets a déjà été développée par plusieurs recherches d'orientation critique. Celles-ci supportent l'hypothèse qu'au-delà de l'amélioration du système visée par ces politiques, ces dernières participent d'une nouvelle régulation normative des systèmes éducatifs. En d'autres termes « les nouvelles politiques éducatives » (Mons, 2004, 2008), et les SABP mettent en acte un nouveau paradigme au travers duquel sont « pensés » les systèmes d'éducation (Maroy, 2009).

6.5 Des effets normatifs et systémiques qui méritent d'être questionnés

La littérature qui s'inscrit dans une approche critique pour questionner l'évolution des modes de régulation nous amène en effet à penser les changements qui opèrent au sein de systèmes éducatifs au-delà de leur apparente technicité. Deux aspects clefs et interreliés nous paraissent ici essentiels pour analyser le développement des SABP dans les systèmes éducatifs : le mouvement de rationalisation de l'éducation qui opère par la mesure et la quantification.

6.5.1 La rationalisation du pilotage des systèmes éducatifs : des liens aux systèmes de production

Tel que nombre d'auteurs l'ont pointé, le développement de la régulation des systèmes éducatifs par leurs résultats s'inscrit dans un mouvement de rationalisation qui touche le secteur public, l'éducation et la société au sens large (Ball, 2003). Cette rationalité technique (Court, 2004) qui peut être aisément éclairée à la lumière des préceptes du *New Public Management*, participe de la la définition des systèmes d'éducation en termes performatifs :

“What do I mean by performativity? Performativity is a technology, a culture and a mode of regulation that employs judgements, comparisons and displays as means of incentive, control, attrition and change based on rewards and sanctions (both material and symbolic). The performances (of individual subjects or organizations) serve as measures of productivity or output, or displays of ‘quality’, or ‘moments’ of promotion or inspection. As such they stand for, encapsulate or represent the worth, quality or value of an individual or organization within a field of judgement” (Ball, 2003, p. 216)

De façon croissante, les systèmes d'éducation sont considérés comme des « systèmes de production » (Maroy, 2009) permettant l'acquisition de savoirs et de compétences utiles pour l'insertion des individus sur le marché de l'économie de la connaissance. Ce que Ball nomme « nouvel ordre moral » (Ball, 1997 ; 2003) favoriserait ainsi la transformation de l'éducation en un « ensemble de moyens que l'on souhaite efficaces, efficaces, à la disposition des individus, des familles et des groupes sociaux dans leur projet de quête ou de maintien de statuts, sur différents marchés » (Lessard, 2002, p. 11).

Les politiques de régulation par les résultats et les SABP, malgré leurs formes différenciées reposent communément sur l'idée d'une standardisation des objectifs à atteindre et, dans une moindre mesure, des moyens pour y parvenir. Comme le discute Maroy l'idée d'une standardisation possible des résultats et des moyens « a été largement utilisée pour améliorer la qualité dans l'industrie » (Maroy, 2009, p. 216). Elle n'est toutefois pas sans poser de question lorsqu'il s'agit du champ éducatif. La montée d'un vocable industriel et économique appliquée à l'éducation est en effet frappante et se retrouve dans nombre productions, au même titre que dans ce présent travail : rendement, efficacité, efficience, qualité, performance.

Au-delà du registre discursif, et des termes dans lesquels sont « pensés » les systèmes d'éducation, les politiques de régulation par les résultats introduisent un ensemble d'instruments permettant un pilotage rationnel du système. Elles reposent sur la fixation de standards, de cibles à atteindre, engagent la mise en place d'instruments de mesure et d'évaluation des résultats des acteurs et des processus organisationnels, et induisent une centration sur les produits de l'éducation allant ainsi de pair avec la standardisation des objectifs à atteindre.

S'il s'agit de soumettre les acteurs à des exigences communes, une standardisation accrue entraîne le risque d'une possible négation de la complexité humaine, de la diversité des individus autant que la multiplicité des contextes qui constituent pourtant le cœur de la relation éducative (Maroy, 2009). La formulation des objectifs du système en termes de résultats et la conception

des systèmes éducatifs comme système de production, tendent par ailleurs à hiérarchiser les fonctions de l'institution scolaire, les fonctions d'éducation et de socialisation étant relayées à des places secondaires au profit de la fonction de distribution⁷¹, plus facilement quantifiable et « standardisable ».

6.5.2 Une rationalisation par la mesure

Cette préoccupation de rationalité va de pair avec un mouvement de quantification de l'éducation qui ne peut être déconnectée des finalités de contrôle et de coordination des systèmes éducatifs, et de leurs acteurs. La mesure de l'éducation fait écho au pouvoir « indubitable » des « chiffres » dans la culture politique moderne, la production de données chiffrées et de statistiques ayant émergé dans nos sociétés modernes libérales démocratiques comme une modalité clef de la production d'une « connaissance nécessaire pour gouverner » (Rose, 1991, p. 676). Au cœur de la quantification et du mouvement de rationalisation de l'éducation s'inscrit la croyance quant à une certaine objectivité des chiffres (des *data*) qui devraient permettre de dresser un portrait plus net de la réalité et d'appuyer une prise de décision plus objective.

Nikolas Rose (1991), dans un brillant essai, met en dialogue différentes approches et conceptions de la quantification de l'action publique et de l'usage des statistiques comme instrument de gouvernance. Ces approches s'opposent par une vision critique axée sur les potentialités des statistiques et des *data*. Elles se rejoignent cependant en ce qu'elles inscrivent leurs analyses dans des préoccupations en termes de rapports et d'exercice du pouvoir. Ces questions, nous semble-t-il, structurent finalement le débat qui entoure le développement de la régulation par les résultats dans les systèmes éducatifs.

La quantification et la mesure de l'éducation s'inscrivent ainsi dans une volonté de contrôle du système et de ses agents, et impliquent dès lors des rapports de pouvoir. Par la mise en œuvre de dispositifs visant à mesurer et à suivre les résultats des acteurs, il s'agit de rendre le travail des éducateurs plus visible et de procéder à un cadrage plus serré de leurs actions. Au cœur du renforcement du contrôle social des enseignants et des établissements scolaire (Rose, 1991) ce

⁷¹ Comme nous l'avons précisé dans notre cadrage théorique, la fonction de distribution fait référence, selon Duret et Martucelli (1966 cités dans Crahay, 2012, p. 25) à l'attribution de qualifications professionnelles et scolaires possédant une certaine utilité sociale.

qui est en jeu est donc très centralement la marge de manœuvre des professionnels du champ ainsi qu'une restructuration des rapports sociaux et des rôles des acteurs du champ.

Au total, ce qui se joue dans les tendances et les transformations actuelles des systèmes éducatifs est une conception renouvelée de l'éducation et de l'institution scolaire. Il s'agirait toutefois d'observer dans quelle mesure ce nouveau référentiel d'action publique, qui participe de la définition des systèmes éducatifs en termes performatifs (Ball, 2003), est traduit dans les orientations de politiques éducatives dans différents contextes politiques, idéologiques et culturels, les instruments d'action publique qu'elles promeuvent et les théories de la régulation dont ces derniers sont porteurs (voir à ce sujet (Maroy, Pons & Dupuy, 2017).

6.6 Systèmes d'*accountability* basés sur la performance et principes de justice : des conceptions sous-jacentes à expliciter

Pour prolonger la réflexion sur la dimension normative des politiques éducatives, nous souhaitons, dans cette sous-section, engager une réflexion critique et normative sur les conceptions de l'équité au cœur des SABP. Nous souhaiterions ainsi proposer une perspective alternative qui pose la question des conceptions de justice qui sous-tendent les politiques de gouvernance par les résultats et, à travers elle, discuter les référents normatifs au prisme desquels est pensée la justice scolaire.

6.6.1 Justice scolaire : entre idéal d'égalité des chances et idéal d'égalité de résultats

Si notre travail visait à questionner les effets des SABP sur l'équité scolaire, il nous faut par ailleurs mettre en dialogue les objectifs politiques et opérationnels qui entourent la mise en œuvre de ces derniers avec différentes figures de l'égalité. Comme nous l'avons vu dans notre cadrage théorique, la question de l'équité scolaire ne peut-être appréhendée qu'au travers de différentes conceptions de justice qui permettent dès lors de définir quelles sont les inégalités « justes » et les inégalités « injustes » au sein des systèmes scolaires.

Si l'on s'intéresse aux SABP, certains auteurs ont émis l'hypothèse que les politiques de régulation par les résultats constituaient une alternative au modèle d'égalité des chances (Duru-Bellat, 2002 ; Dupriez, Oriane & Verhoeven, 2008). En soumettant les systèmes éducatifs à une obligation de résultat, la visée de ces politiques serait une égalisation des résultats des élèves à

l'issue d'une période de formation donnée, et non le fait de donner, en amont, les mêmes chances de résultats aux élèves (au-delà de leur origine sociale).

L'idéal d'égalité de résultats reposerait ainsi sur une conception égalitariste stricte de la justice scolaire qui implique que les inégalités sont par définition intolérables, « qu'une société juste c'est une société sans inégalités » (Dupriez, Oriane, & Verhoven, 2008, p. 109). En ces termes, il revient à l'institution scolaire de compenser le poids des inégalités qui lui pré-existent et d'assurer, pour tous les élèves, des niveaux similaires d'acquisitions (Dupriez et al., 2008). L'objectif est donc plus ambitieux que dans un idéal d'égalité des chances de résultats où il s'agit de compenser le poids de différents facteurs (géographiques, socioéconomiques) pour offrir aux élèves les mêmes chances de réussite.

Est-ce à dire que les politiques de régulation par les résultats et les SABP proposent réellement une alternative au modèle dominant d'égalité des chances, modèle qui tolère de fait les inégalités de résultats, ces dernières étant considérées comme le reflet du mérite et des capacités des élèves?

Il faut ici se rapporter aux objectifs opérationnels articulés aux SABP pour pouvoir répondre à cette question. La réponse diffère en effet en fonction de ce qui est poursuivi : s'agit-il de l'égalité de résultats? De l'égalité des acquis de base? Dans le premier cas, l'impératif a trait à un niveau d'acquisitions sensiblement égal entre élèves, au-delà d'un ensemble de facteurs qui s'avèrent déterminants dans leur réussite. Ce qui est visé est nettement plus ambitieux et les inégalités de résultats sont dès lors intolérables et doivent être combattues.

Dans le second cas, c'est-à-dire dans une perspective d'égalité des acquis de base, les différences de performance à l'issue d'une période de formation sont considérées comme « justes », sous réserve que les élèves parviennent à un niveau commun d'acquisition. On peut supposer que la méritocratie reste ici finalement le fondement normatif qui prévaut. Les inégalités de résultats au sein de l'institution scolaire sont ainsi tolérées, si tant est que chacun soit outillé d'un bagage minimal pour s'insérer dans la société.

6.6.2 Pour une vision plus ambitieuse de la justice scolaire

Au total, il nous semble que les politiques éducatives doivent se doter d'objectifs plus ambitieux. La poursuite de l'égalité des acquis de base, au moins dans la scolarité obligatoire, apparaît comme un objectif politique nécessaire, mais insuffisant. Les visées d'égalisation de résultats doivent par ailleurs être assorties de mesures volontaristes de lutte contre les inégalités d'origines sociales et économiques, mais aussi liées à des conditions de scolarisation sensiblement inégales liées, par exemple, aux phénomènes de ségrégation sociale et scolaire.

S'il s'agit d'agir sur les inégalités scolaires, il semble que les systèmes d'incitatifs (tel que le suggère notre travail) et de sanctions graduées au cœur de certains SABP (dans le contexte des États-Unis) apparaissent clairement inefficaces, voire contre-productifs, dans cette lutte contre les inégalités. Les SABP à enjeux forts entraîneraient, qui plus est, des effets négatifs voire pervers sur les publics les plus fragiles (Wolf & Janssens, 2007; Finnigan & Gross, 2007; Mintrop & Sunderman, 2009 ; Figlio & Loeb, 2011). On peut alors se demander si ce type de dispositifs, et plus particulièrement les systèmes de sanctions graduées, ne supportent pas une logique qui revient à donner moins à ceux qui en ont déjà le moins.

Par ailleurs, la régulation fédérale établie aux États-Unis avec *No Child Left Behind*, en posant comme condition à la pérennité des financements compensatoires octroyés aux établissements en difficultés l'efficacité des établissements scolaires, envoie un message clair : la politique de financement compensatoire n'a de raison d'être que si et seulement si les résultats escomptés sont effectivement au rendez-vous après une période de temps déterminée. On peut dès lors se demander dans quelle mesure l'égalisation de résultats devient possible sans un soutien fort de l'État?

En outre, certains facteurs structurels (structure sociale des systèmes d'éducation, stratification du système), telles que la littérature et notre recherche l'ont montré, s'avèrent clefs dans la capacité des systèmes éducatifs à réduire ou à amplifier les inégalités scolaires (et les inégalités scolaires d'origine sociale). Il semble donc pertinent de questionner l'organisation des systèmes scolaires et ses effets sur l'équité scolaire. Si plusieurs travaux ont par ailleurs montré que les pratiques éducatives et les pratiques de gestion mises en œuvre au sein des établissements scolaires sont des facteurs vecteurs d'efficacité et d'équité scolaires (voir entre autres Välijärvi et al. (2002)) la pertinence de mettre en place des mesures visant à soutenir le développement des

professionnels du champ n'apparaît que d'une plus grande acuité.

6.7 Une autre voie possible?

S'il semble difficile d'infléchir les tendances en cours et de contrer la popularité des politiques d'*accountability*, il nous paraît nécessaire, pour clore la discussion de ce travail de recherche, de discuter d'une autre voie possible pour penser ces politiques.

Il ne s'agit pas de nier certains éléments fondamentaux sur lesquels repose l'acception de la notion d'*accountability* et qui nous semblent essentiels avec, au premier chef, l'idée de responsabilité. Il est évident que l'institution scolaire a une responsabilité majeure dans l'instruction et l'éducation des enfants qui lui sont confiés. Il est donc incontestable que l'institution scolaire doit rendre des comptes sur sa capacité d'instruire et d'éduquer. Si l'éducation est un droit reconnu dans la Déclaration universelle des droits de l'Homme, c'est aussi un projet politique, social et sociétal, un pilier fondamental de nos sociétés. C'est un instrument puissant qui ne saurait faire l'économie d'une réflexion en profondeur sur les finalités qui lui sont attachées, les modalités et moyens associés à leur réalisation.

Or, si les objectifs politiques d'efficacité et d'équité des systèmes éducatifs associés au développement de la gouvernance par les résultats en éducation semblent difficilement contestables, une voie alternative pour penser les politiques d'*accountability* et les SABP semble nécessaire.

6.7.1 La centration sur les « produits » : vers des objectifs plus larges

Premièrement, il ne s'agit pas d'occulter la nécessité pour les systèmes éducatifs de se doter d'objectifs à atteindre et de poser des cadres pour soutenir l'action des éducateurs. Cependant, la tendance à la centration sur les produits de l'éducation et sur les fondamentaux, en d'autres termes, la littératie et la numératie, au cœur des tendances globales en matière de politiques éducatives et des SABP (Sahlberg, 2007, 2010, 2011) sont largement discutables.

Si les systèmes d'éducation doivent se doter de standards en termes de connaissances et de compétences à atteindre, les objectifs du système ne peuvent être formulés qu'en termes d'acquisitions scolaires et d'objectifs cognitifs au détriment d'objectifs plus larges (socialisation, compétences sociales, valeurs morales). L'institution scolaire possède aussi une fonction

d'éducation et de socialisation qui doit être mise en l'avant dans les objectifs des politiques et des systèmes éducatifs. La fonction éducative renvoie ici au développement de la personne et de ses capacités de réflexivité, celle de socialisation entre autres à l'intégration de valeurs communes (Dupriez et al., 2008). Il en va finalement du rôle de l'école comme vecteur de cohésion sociale qui doit être réaffirmé.

Ainsi, une approche flexible, valorisant des standards plus « lâches » (Sahlberg, 2007; 2011; Maroy & Voisin, 2017) intégrant des objectifs cognitifs et non cognitifs, qui respecte le rythme d'apprentissage des élèves tout en poursuivant des niveaux d'exigences élevés, semble constituer une alternative intéressante pour pallier les critiques et les effets négatifs des approches agressives par les standards de performance décriées par de nombreux auteurs.

Aussi, si les politiques d'*accountability* visent à réduire l'opacité du système ainsi qu'à réduire le découplage entre les différents niveaux d'action, nous avons vu le rôle central joué par les dispositifs d'évaluation (des établissements scolaires, des acquis des élèves). Ceux-ci permettent en effet d'une part de renforcer le contrôle organisationnel et social sur les acteurs (institutions et professionnels), mais aussi de recueillir de l'information quant à la nécessité de pourvoir à d'éventuels ajustements en termes de processus de gestion ou de processus éducatifs. Il nous semble, à l'instar de Demailly (2001), que les dispositifs d'évaluation peuvent donner lieu et place à une réflexion qui peut s'avérer porteuse à partir du moment où ces dispositifs sont envisagés (par les enseignants et les acteurs impliqués) comme des potentialités d'innovation et de développement professionnel.

6.7.2 Enjeux et responsabilisation des acteurs : vers une responsabilité collective

Les systèmes d'incitatifs et de sanctions sur lesquels reposent certains types d'*accountability*, de même que le principe de responsabilité individuelle et l'imputabilité des acteurs du champ éducatif, en premier lieu les enseignants, doivent aussi être questionnés. Pour Sahlberg (2007 ; 2010; 2011) un système d'*accountability* « intelligente » repose principalement sur un niveau de confiance élevé à l'égard du professionnalisme enseignant (et des chefs d'établissements), de leur jugement professionnel et de leur capacité à mettre en place des conditions d'apprentissage optimales, ainsi qu'à mobiliser les pratiques pédagogiques les plus adaptées aux besoins des élèves. Si les enseignants ont effectivement un rôle majeur à jouer dans les apprentissages des

élèves et dans le développement de leurs compétences, ils doivent pouvoir s'appuyer sur une formation initiale de qualité et sur des dispositifs de soutien leur permettant de développer leurs pratiques et d'engager une réflexion sur ces dernières.

Qui plus est, tel que nous l'avons mentionné, nous considérons à l'instar d'autres auteurs que les structures (sociales, organisationnelles) déterminent en partie les comportements des acteurs et influent sur leur marge de manœuvre. Le travail enseignant s'inscrit dans un système au sens de Crozier et Friedberg (1977), dans un cadre structurel comportant des contraintes organisationnelles. Le travail enseignant et la relation d'apprentissage se construisent par ailleurs au quotidien au travers d'un ensemble d'interactions avec les élèves, mais aussi avec d'autres acteurs du champ éducatif (collègues, personnel de soutien, hiérarchie) qui doivent être prises en compte pour discuter du principe de responsabilité. Ici, la question de savoir s'il est possible de séparer « ce qui est imputable à l'action de l'organisation (e.g scolaire) et de ces agents, et ce qui relève d'un effet de contexte et d'environnement qui lui échappe » (Maroy, 2009, p. 214) s'avère un défi important qui rend questionnable, de notre point de vue, l'imputabilité des acteurs et, au premier chef, des enseignants.

Ainsi, le rôle et l'implication de l'ensemble des acteurs dans la mise en œuvre d'un projet commun d'éducation doivent être affirmés. Par ailleurs le principe de responsabilité ne peut être envisagé, de notre point de vue, qu'en termes collectifs. Si les enseignants doivent être en mesure de mettre en place les stratégies pertinentes pour favoriser les apprentissages des élèves, les autorités administratives doivent fournir les conditions et moyens nécessaires à un enseignement de qualité, les systèmes d'éducation doivent aussi être questionnés à la lumière de facteurs structurels et organisationnels.

Finalement la réflexion sur les systèmes d'éducation et l'école ne peut être déconnectée d'une réalité sociale et structurelle plus large, d'autant plus si l'objectif est de les interpeller à la lumière de critères de justice sociale. Si l'on attend des systèmes d'éducation la réussite de tous les élèves et la réduction des inégalités scolaires d'origine socioculturelle et économique, la société au sens large doit aussi s'interroger sur le caractère inacceptable des inégalités qui existent en son sein.

CONCLUSION

Notre recherche constitue l'une des rares études quantitatives comparatives internationales ayant permis d'une part d'analyser les formes que revêtent les SABP dans les systèmes d'éducation d'Europe et du Canada et, d'autre part, d'interroger leurs effets sur les degrés d'efficacité et d'équité scolaires de ces systèmes d'éducation.

En nous focalisant sur les instruments mis en œuvre pour réguler le fonctionnement des établissements scolaires sur base de leurs résultats, nous avons pu tout d'abord, dans un premier volet de recherche, caractériser les SABP dans ces systèmes d'éducation. Cette entrée d'analyse par les instruments renseignés par notre base de données, ancrée dans un cadrage conceptuel mobilisant principalement les notions de régulation et de modes de régulation institutionnelle, nous a permis de mettre en exergue les logiques instrumentales au cœur des SABP et d'élaborer sur cette base l'une des premières typologies empiriques permettant de caractériser les SABP dans les systèmes éducatifs étudiés.

Ainsi, la démarche que nous avons adoptée apporte plusieurs contributions substantielles à l'étude de l'objet. Nous avons en effet souhaité prendre en compte les instruments de régulation à l'œuvre au sein des établissements afin d'observer ces modes de régulation à l'œuvre et leurs effets sur l'efficacité et l'équité scolaires. Ainsi, nous ne nous sommes pas contentés de prendre en compte certains instruments pour tester une théorie de l'action spécifique, mais avons souhaité observer dans quelle mesure les modes de régulation mis en œuvre au sein des établissements scolaires s'articulent pour mettre en œuvre des logiques instrumentales distinctes dès lors au cœur de SABP. De même, nous avons étudié la combinaison de ceux-ci au sein de types empiriques distincts permettant de regrouper ou de différencier les systèmes éducatifs d'Europe et du Canada.

Trois logiques instrumentales émergent ainsi de nos analyses : une logique d'évaluation et de régulation selon la performance des acteurs, une logique reposant sur des incitatifs directs pour les enseignants, une logique de comparaison des performances des établissements scolaires où des dispositifs visant à agir sur les processus internes aux établissements sont par ailleurs prévus (intervention d'experts externes, participation des enseignants à des activités de développement professionnel). Sous-tendues par des instruments distincts, ces logiques s'articulent au sein de

types empiriques de SABP qui permettent de différencier les systèmes d'éducation d'Europe et du Canada : 1) une *accountability* de faible intensité et peu instrumentée, 2) une *accountability* axée principalement sur l'évaluation et le suivi de la performance, 3) une *accountability* articulant suivi de la performance, comparaison externe et amélioration interne des résultats des établissements scolaires ; 4) une *accountability* de forte intensité privilégiant la mise en place d'incitatifs directs pour les enseignants.

Dans un second volet de recherche, nous avons par ailleurs analysé les effets de ces logiques et types empiriques de SABP sur l'efficacité et l'équité scolaires des systèmes éducatifs d'Europe et du Canada. Les résultats des analyses menées nous amènent toutefois à dresser le constat d'un impact relativement ténu des SABP sur l'efficacité et l'équité scolaires. En effet, nos analyses suggèrent que les effets des SABP sur les degrés d'efficacité et d'équité scolaires des systèmes éducatifs d'Europe et du Canada sont, dans la majorité des cas, faibles et s'annulent le plus souvent lorsque des facteurs-clefs sont pris en compte (dépenses publiques en éducation, variation du niveau socioéconomique des élèves, stratification horizontale et verticale des systèmes scolaires). Si la mise en place d'incitatifs pour le corps enseignant et le type de SABP privilégiant ce type de dispositifs apparaissent significativement liés à des degrés d'efficacité scolaire plus faibles, et d'inégalités de résultats et des acquis de base plus élevés, nos résultats invitent finalement à dresser un constat plus que mitigé quant à la capacité des SABP à tenir leurs promesses d'efficacité et d'équité scolaires. Il ressort ainsi de nos analyses que les SABP ne permettent pas, tout du moins dans les contextes étudiés, de rendre les systèmes éducatifs plus efficaces et plus équitables. Ce constat rejoint ainsi les conclusions de plusieurs auteurs qui remettent finalement en question les présupposés qui entourent la mise en œuvre des SABP dans le champ éducatif.

Il semble dès lors nécessaire de questionner la rhétorique qui entoure le développement des SABP dans le champ éducatif, et plus largement de la nouvelle régulation des systèmes éducatifs par leurs résultats. Puisqu'il nous semble que les tendances actuelles à l'œuvre dans le champ éducatif doivent être interpellées (focalisation sur des apprentissages strictement cognitifs et sur les fondamentaux [littératie, numératie] au détriment d'objectifs plus larges [ayant trait à la socialisation des individus, compétences sociales, valeurs morales], accent mis sur l'imputabilité des acteurs [plus particulièrement des enseignants] et sur la responsabilité individuelle au détriment d'une responsabilité collective dans la mise en œuvre d'un projet commun

d'éducation), nous avons proposer des pistes en termes plus normatifs pour penser le développement des SABP dans les systèmes d'éducation.

Par ailleurs, nous avons mis en exergue la nécessité pour les systèmes d'éducation de se doter d'objectifs plus ambitieux dans leur quête de justice scolaire. Il nous semble en effet que l'égalité des acquis de base, visant à permettre à chacun d'acquérir un ensemble de connaissances et compétences nécessaires à la participation active à la société, apparaît comme un objectif nécessaire, mais insuffisant. Cette quête de justice scolaire ne peut, de notre point de vue, par ailleurs, faire l'économie de mesures volontaristes de lutte contre les inégalités au sein et à l'extérieur de l'institution scolaire.

Si la mise en œuvre des SABP s'inscrit dans une préoccupation d'équité scolaire, notre thèse, que les SABP ne parviennent pas, à eux seuls, à réduire le lien entre inégalités sociales et inégalités scolaires ni, plus généralement, l'écart de résultats entre élèves. D'autres facteurs, comme la stratification horizontale et verticale (autrement nommée différenciation verticale et horizontale [Mons, 2004]), s'avèrent clefs dans l'amplification des inégalités scolaires et doivent donc être interrogés.

En conclusion, nous reprenons ici, à l'instar des travaux d'orientation critique, la question déjà évoquée des effets normatifs des SABP. Ces effets normatifs ont, entre autres, trait à l'émergence d'un nouveau référentiel d'action en éducation participant de la définition des systèmes d'éducation comme des systèmes de production (Maroy, 2009) et d'une obligation de résultats en éducation (Lessard & Meirieu, 2004 ; Demailly, 2001) dont les implications doivent être mises au jour. Si on accepte le postulat de Muller selon lequel les politiques publiques « constituent des espaces au sein desquels les acteurs concernés vont construire et exprimer 'un rapport au monde' qui renvoie à la manière dont ils perçoivent le réel, leur place dans le monde et ce que le monde devrait être » (Muller, 2000, p. 196), cela nous amène à interroger les politiques d'*accountability* dans des termes qui dépassent la question, bien que cruciale, de leurs effets en termes d'équité et d'efficacité scolaires.

LIMITES ET PERSPECTIVES DE RECHERCHE

Si notre thèse doctorale contribue certainement au champ scientifique de l'analyse de politiques éducatives, elle présente certaines limites qui ouvrent cependant des perspectives de recherche. Nous les discutons dans cette dernière sous-section.

Premièrement, notre recherche doctorale s'est appuyée sur les données du PISA collectées via une approche extensive par questionnaire, nous amenant à travailler sur des données autorapportées pour caractériser les SABP dans les systèmes d'éducation. Bien que cette base de données constitue l'une des sources d'informations comparatives sur les systèmes d'éducation parmi les plus riches, certaines variables pertinentes n'ont pas pu être prises en compte dans l'analyse, telles que la présence d'examens externes standardisés (Carnoy & Loeb, 2002; Mons, 2009; 2013; Yerly, 2014). Bien que nous ayons mis en exergue dans notre revue de littérature le fait que la grande majorité des systèmes éducatifs étudiés avaient effectivement en place ce genre de dispositifs, la fréquence de ces évaluations, mais aussi les niveaux d'enseignement concernés devraient être pris en compte et intégrés à l'analyse. Ces informations devraient dès lors être recueillies systématiquement par les grandes enquêtes internationales du fait de leur rôle central et croissant au sein des systèmes éducatifs.

Deuxièmement, comme nous l'avons précisé, la typologie empirique des SABP que nous avons proposée dans ce travail de thèse revêt un caractère exploratoire. Elle est dès lors fortement liée aux caractéristiques de l'échantillon à partir duquel la recherche a été menée. Elle devrait alors être affinée dans des recherches à venir. Ici premièrement, notre recherche a montré des pistes qui devraient être explorées pour typifier les SABP dans les contextes étudiés. Si nous avons observé la formation de sous groupes au sein des *clusters*, ces sous-groupes devraient être davantage pris en compte pour affiner la typologie, ce que nous n'avons pu faire puisque nous étions contrainte par les méthodes d'analyse subséquentes mobilisées. Qui plus est la typologie mériterait d'être étudiée via les dernières données du PISA disponibles (PISA 2015) pour observer l'évolution, le cas échéant, des systèmes d'éducation quant à l'usage des instruments étudiés dans la présente thèse. En ce qui a trait à l'échantillon sur lequel porte la recherche, l'ensemble des *Lands* allemands devrait être pris en compte de même que les cantons suisses. Par ailleurs, les systèmes éducatifs des États-Unis devraient être intégrés à l'analyse car ils constituent un point de référence pertinent du fait des nombreuses recherches menées dans ces

contextes. Il faut cependant que les données soient disponibles et accessibles pour ce faire. Ceci renvoie à la nécessité pour l'OCDE, et les enquêtes PISA, de mettre à disposition les données désagrégées par systèmes d'éducation.

Troisièmement, notre étude n'a pas permis de procéder à une analyse plus fine des dispositifs d'*accountability* renseignés par notre base de données, de leurs formes et de leurs modalités concrètes de mise en oeuvre. Or, tel que la littérature l'a montré, des dispositifs *a priori* relativement similaires peuvent, dans les faits, revêtir des modalités concrètes d'application sensiblement différentes d'un contexte à l'autre (Maroy, 2013). S'agissant par exemple des dispositifs d'évaluation externe des établissements, il serait intéressant de voir quels sont les éléments pris en compte dans l'évaluation des établissements (processus, produits)? Quels sont les acteurs « évaluateurs »? Selon quelle temporalité les évaluations sont-elles réalisées?

En outre, quatrième point, puisque les instruments d'action publique structurent par ailleurs les rapports sociaux entre acteurs (Lascoumes & Legales, 2004 ; Lascoumes & Simard, 2011), d'autres recherches devraient être menées *in situ* pour développer une analyse plus fine des dispositifs renseignés par notre base de données, de leurs usages, des enjeux qu'ils revêtent. Qui plus est, il nous faudrait observer dans quelle mesure les instruments d'*accountability* interagissent et influent sur les pratiques à l'œuvre au sein des établissements scolaires.

Ainsi si les analyses mobilisées dans notre recherche doctorale nous ont permis de mettre en exergue des grandes tendances en ce qui a trait aux relations entre SABP, efficacité et équité scolaires on peut par ailleurs faire l'hypothèse, appuyée sur la littérature, que d'autres variables interviennent par ailleurs dans cette relation. Si nous nous sommes focalisée sur des facteurs structurels au niveau des systèmes éducatifs (stratification du système), d'autres facteurs intervenant à différents niveaux d'action (des établissements – caractéristiques des établissements, pratiques de gestion [etc.] – et des classes – caractéristiques des élèves, pratiques éducatives [etc.]) mériteraient d'être pris en compte. Ici, il apparaît donc que l'usage de modèles multi-niveaux permettrait d'une part, un meilleur contrôle de l'effet de ces facteurs, mais aussi d'observer les interactions entre les modes de régulation implantés au niveau des établissements scolaires, les pratiques éducatives et les pratiques de gestion à l'œuvre. Nous pourrions ainsi mieux saisir les effets des SABP et de leurs instruments sur ces dimensions, qui demeurent un autre enjeu majeur au cœur de la mise en œuvre des SABP.

Dans la même veine, des recherches mixtes doivent par ailleurs être développées pour pouvoir appréhender les processus à l'œuvre dans la mise en œuvre des SABP. Ce type de recherche permettrait en effet de faire le pont entre les grandes tendances que des recherches comme la nôtre, à grande échelle, permettent de saisir, et les processus à l'œuvre (processus de médiation à l'œuvre dans la mise en œuvre des politiques éducatives) que les recherches qualitatives, par un grain plus fin d'analyse, offrent la possibilité d'appréhender.

Enfin, l'analyse des politiques éducatives ne peut faire l'économie d'une réflexion critique sur les implications systémiques et normatives des dispositifs et instruments mis en œuvre par ces dernières. Ces aspects sont fortement complémentaires et essentiels à l'analyse des politiques éducatives et de leur implémentation au sein des systèmes d'éducation.

BIBLIOGRAPHIE

- Amrein, A., & Berliner, D. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18).
- Attali, A., & Bressoux, P. (2002). *L'évaluation des pratiques éducatives dans les premier et second degrés*. Retrieved from Paris
- Baillargeon, J. (2003). L'analyse en composantes principales. [Document inédit disponible en ligne]. Retrieved from <http://www.uqtr.ca/cours/srp-6020/acp/acp.pdf>
- Ball, S. J. (1997). Policy sociology and critical social research: a personal review of recent education policy and policy research. *British educational research journal*, 23(3), 257-274.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215-228.
- Barbana, S., Dumay, X., & Dupriez, V. (2014). Perceptions et usages des instruments d'accountability. Enquête exploratoire dans l'enseignement secondaire en Belgique francophone. *Education Comparée*, 12, 21-44.
- Barnard, C. I. (1938). The economy of incentives. *Classics of organization theory*, 93-102.
- Barrère, A. (2010). Ce que fait l'évaluation aux établissements scolaires. Une année dans un collège d'« éducation prioritaire ». *Ethnologie française*, 40(1), 141-149.
- Barroso da Costa, C. (2014). *L'engagement professionnel chez les nouveaux enseignants et la satisfaction des gestionnaires d'école à l'égard du travail effectué par les enseignants novices*. (Thèse de doctorat), Université de Montréal, Montreal.
- Barroso, J. (2000). Autonomie et modes de régulation locale dans le système éducatif. *Revue française de pédagogie*, 130, 57-71.
- Battaglia, M., & Collas, A. (2013). La France championne des inégalités. *Le monde* (3 décembre).
- Behrens, M. (2006). *Aanalyse de la littérature critique sur le développement, l'usage et l'implémentation des standards dans un système éducatif*. Retrieved from
- Béland, S., Cousineau, D., & Loye, N. (2016). Les dix commandements du nouvel homo statisticus. *Mc Gill Journal of Education*, 51(2), 947-960.
- Ben Jaafar, S., & Earl, L. (2008). Comparing Performance-Based Accountability models : a canadian example. *Canadian Journal of Education*, 31(3), 697-726.
- Berényi, E., Bajomi, I., & Neumann, E. (2013). Une expérience hongroise visant à gouverner par les nombres – genèse et évolution du Système National de Mesure des Compétences. In C. Maroy & Y. Dutercq (Eds.), *L'école à l'épreuve de la performance. Les politiques de régulation par les résultats*. Bruxelles: De Boeck.
- Bezes, P. (2005). Le renouveau du contrôle des bureaucraties. *Informations sociales*, 126(6), 26-37.
- Bishop, J., Mane, F., Bishop, M., & Moriarty, J. (2001). *The role of end-of course exam and minimum competency exams in Standards-Based Reforms*. Retrieved from Ithaca, New York:

- Blashfield, R. K., & Aldenderfer, M. S. (1988). The methods and problems of cluster analysis *Handbook of multivariate experimental psychology* (pp. 447-473): Springer.
- Boudon, R. (1973). L'Inégalité des chances. *La mobilité sociale dans les sociétés industrielles*, Paris, Armand Colin.
- Boudon, R. (2004). Théorie du choix rationnel ou individualisme méthodologique. *Revue du MAUSS*, 2(24), 281-309.
- Bourdieu, P. (1966). L'école conservatrice. Les inégalités devant l'école et devant la culture. *Revue française de pédagogie*, 7(3), 325-347.
- Bourdieu, P., & Passeron, J.-C. (1964). *Les héritiers. Les étudiants et la culture*. . Paris: Les éditions de minuit.
- Bouroche, J. M., & Saporta, G. (1980). *L'analyse des données*. Paris: Presses Universitaires de France.
- Bourque, J., Poulin, N., & Cleaver, A. F. (2006). Évaluation de l'utilisation et de la présentation des résultats d'analyses factorielles et d'analyses en composantes principales en éducation. *Revue des sciences de l'éducation*, 32(2), 325-344.
- Bovens, M. (2006). Analysing and assessing public accountability. A conceptual framework. *European Governance Papers (EUROGOV)*, N. C-06-01.
- Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, 13(4), 447-468.
- Bovens, M. (2010). Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. *West European Politics*, 33(5), 946-967.
- Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211-252.
- Boyd, W. (2003). *American educational on trial change and challenges*. Chicago: The University of Chicago Press.
- Brassard, A. (2009). Le projet de loi 88 : ses implications sur l'autonomie de l'établissement et sur les relations entre les échelons du système. *Le Point en administration de l'éducation*, 11(3), 11-14.
- Brassard, A., Lusignan, J., & Lessard, C. (2002). *Étude de cas sur l'expérience des changements en éducation dans trois commissions scolaires du Québec*. Montréal: Université de Montréal, FSE, LABRIPROF-CRIFPE.
- Bressoux, P. (2010). *Modélisation statistique appliquée aux sciences sociales: De boeck Brussels*.
- Broadfoot, P. (2000). Un nouveau mode de régulation dans un système décentralisé: l'État évaluateur. *Revue française de pédagogie*, 130(1), 43-55.
- Bruns, B., Filmer, D., & Patrinos, H. A. (2011). *Making schools work: New evidence on accountability reforms*: World Bank Publications.
- Campbell, J. L. (2002). Pour convaincre les sceptiques : à propos des idées et des critiques de la théorie du choix rationnel. *Sociologie et sociétés*, 34(1), 35-50.

- Carnoy, M., Elmore, R., & Siskin, L. (2003). *The new accountability : High schools and high-stakes testing* New-York: Routledge.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes ? A cross-state analysis. *Educational Evaluation & Policy Analysis*, 24(4), 305-331.
- Cattell, R. B. (1978). *The scientific use of factor analysis*. New-York: Plenum.
- Chan, A., Fisher, D., & Rubenson, K. (2007). *The evolution of professionalism : Educational policy in Provinces and Territories of Canada*. . Vancouver: University of British Columbia Press.
- Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, 93(9-10), 1045-1057.
- Chubb, J. E., & Moe, T. M. (1988). Politics, Markets, and the Organization of Schools. *The American Political Science Review*, 82(4), 1065-1087.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, Markets and America's Schools*. Washington, D.C.: The Brookings Institution.
- Clark, P., & Wilson, J. (1961). Incentive Systems: A Theory of Organizations. *Administrative Science Quarterly*, 6(2), 129-166.
- Coburn, E. C. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211-244.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Washington: US Government Printing Office
- Commaille, J., & Jobert, B. (1998). *Les métamorphoses de la régulation politique*. Paris: Librairie Générale de Droit et de Jurisprudence.
- No Child Left Behind Act of 2001 - Public Law 107-110, United States Congress, (2002).
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational research methods*, 6(2), 147-168.
- Cosnefroy, O., & Rocher, T. (2004). Le redoublement au cours de la scolarité obligatoire: nouvelles analyses, mêmes constats. *Éducation & formations*, 70, 73-82.
- Council, C. E. S. (2013). *Education indicators in Canada: An international perspective*: Ottawa, Canada: Canadian Education Statistics Council.
- Crahay, M. (2004). Peut-on conclure à propos des effets du redoublement ? *Revue française de pédagogie*, 148, 11-23.
- Crahay, M. (Ed.) (2012). *L'école peut-elle être juste et efficace? De l'égalité des chances à l'égalité des acquis*. (2 ed.). Bruxelles: de Boeck.
- Croxford, L., Grek, S., & Shaik, F. J. (2009). Quality assurance and evaluation (QAE) in Scotland: promoting self- evaluation within and beyond the country. *Journal of Education Policy*, 24(2), 179-193.
- Crozier, M., & Friedberg, E. (1977). *L'acteur et le système*. Paris: Le Seuil.

- Cussó, R., & D'Amico, S. (2005). Vers une comparabilité plus normative des statistiques internationales de l'éducation : de l'éducation de masse aux compétences. *Cahiers de la recherche sur l'éducation et les savoirs, Hors-Série n°1 Pouvoirs et mesure en éducation*, 21-47.
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *The Teacher College Record*, 106(6), 1047-1085.
- Darling-Hammond, L. (2006). *Standards, assessments, and educational policy: In pursuit of genuine accountability*. Princeton, NJ: Educational Testing Service.
- de Landsheere, G. (1994). *Le pilotage des systèmes d'éducation*. Bruxelles: De Boeck.
- Demailly, L. (2001). *Évaluer les politiques éducatives. Sens, enjeux, pratiques*. Bruxelles: De Boeck
- Demailly, L. (2003). L'obligation de résultats en éducation. *Administration et éducation*, 2, 67-84.
- Demailly, L. (2006). En Europe : l'évaluation contre la crise des systèmes scolaires, l'évaluation en crise. *Education et sociétés*, 1(17), 105-120.
- Dembélé, M., Goulet, S., Lapointe, P., & Deniger, M.-A. (2013). Perspective historique de la construction du modèle québécois de gestion scolaire axée sur les résultats. In C. Maroy (Ed.), *L'école à l'épreuve de la performance : Les politiques de régulation par les résultats : trajectoires nationales, usages locaux*: DeBoeck.
- Demeuse, M., & Baye, A. (2009). Indicateurs d'équité éducative. Une analyse de la ségrégation académique et sociale dans quelques pays européens. *Revue française de pédagogie*, 165, 91-103.
- Dubet, F. (2010). Les inégalités scolaires entre l'amont et l'aval. Organisation scolaire et emprise des diplômes. *Sociologie*, 1(2), 177-197.
- Dubet, F., & Martuccelli, D. (1996). Théories de la socialisation et définitions sociologiques de l'école. *Revue française de sociologie*, 37(4), 511-535.
- Dulude, E., & Dupriez, V. (2014). Les politiques et les pratiques de pilotage par les résultats. Perspectives internationales. *Education Comparée*, 12.
- Dumay, X. (2010). Évaluation et accompagnement des établissements en Europe : diversité et mécanismes d'hybridation. *Education et sociétés*, 2(26), 145-160.
- Dupriez, V., & Dumay, X. (2005). L'égalité des chances à l'école : analyse d'un effet spécifique de la structure scolaire. *Revue française de pédagogie*, 150, 5-17.
- Dupriez, V., & Dumay, X. (2011). Les marchés scolaires : au bénéfice de qui ? *Revue française de pédagogie*, 176, 83-100.
- Dupriez, V., Oriane, J. F., & Verhoven, M. (2008). *De l'école au marché du travail, l'égalité des chances en question*. Berne: Peter Lang.
- Dupriez, V., & Vandenberghe, V. (2008). Les inégalités dans les systèmes scolaires en Europe : état des lieux. In V. Dupriez, J. F. Oriane, & M. Verhoeven (Eds.), *De l'école au marché du travail, l'égalité des chances en question*. Berne: Peter Lang.

- Durkheim, E. (1894). Les règles de la méthode sociologique. *Revue Philosophique de la France et de l'Étranger*, 37, 465-498.
- Duru-Bellat, M. (2002). *Les inégalités sociales à l'école. Genèse et mythes*. Paris: Presses Universitaires de France.
- Duru-Bellat, M., Jarousse, J. P., & Mingat, A. (1993). Les scolarités de la maternelle au lycée: étapes et processus dans la production des inégalités sociales. *Revue française de sociologie*, 34(1), 43-60.
- Duru-Bellat, M., & Meuret, D. (2001). Note de synthèse : Nouvelles formes de régulation dans les systèmes éducatifs étrangers : autonomie et choix des établissements. *Revue française de pédagogie*, 135.
- Duru-Bellat, M., & Mingat, A. (1988). Le déroulement de la scolarité au collège : le contexte « fait des différences ». *Revue française de sociologie*, 29(4), 649-666.
- Duru-Bellat, M., Mons, N., & Suchaut, B. (2004). Caractéristiques des systèmes éducatifs et compétences des jeunes à 15 ans. L'éclairage des comparaisons entre pays. *Les Cahiers De L'IREDU*, 66, 1-158.
- Dutercq, Y., & Cuculou, S. (2013). La performance comme outil de gouvernance : quelles conséquences sur le travail des enseignants de l'école primaire ? . In C. Maroy (Ed.), *L'école à l'épreuve de la performance* (pp. 195-210). Bruxelles: De Boeck.
- Dutercq, Y., & Maroy, C. (2017). *Le travail et le professionnalisme enseignant face aux politiques de responsabilisation*. Bruxelles: De Boeck.
- Dutercq, Y., & Van Zanten, A. (2011). Présentation : l'évolution des modes de régulation de l'action publique en éducation. *Education et sociétés*, 8, 5-10.
- Eurydice. (2008). *Responsabilité et autonomie des enseignants en Europe*. Retrieved from Luxembourg : Publications Office of the European Union:
- Eurydice. (2009). *Les évaluations standardisées des élèves en Europe : objectifs, organisation et utilisation des résultats* Retrieved from Luxembourg : Publications Office of the European Union:
- Eurydice. (2015). *Assuring quality in education: policies and approaches to school evaluation in Europe*. Retrieved from Luxembourg : Publications Office of the European Union:
- Everitt, S. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry*, 126(3), 237-240.
- Everitt, S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis. 5th Edition*. Chichester: Wiley & Sons.
- Falabella, A. (2013). *Accountability policy effects within school markets. A study in three chilan municipalities*. University of London.
- Falabella, A. (2014). The Performing School: The Effects of Market & Accountability Policies. . *Education Policy Analysis Archives*, 22(70). doi:
- Falissard, B. (1996). *Comprendre et utiliser les statistiques dans les sciences du vivant*. Paris: Masson.

- Farvaque, N. (2008). Education, formation, insertion : les approches par les capacités de Sen. In V. Dupriez, J. F. Orienne, & M. Verhoeven (Eds.), *De l'école au marché du travail, l'égalité des chances en question*. Berne: Peter Lang.
- Faure, A. M. (1994). Some methodological problems in comparative politics. *Journal of Theoretical Politics*, 6(3), 307-322.
- Felouzis, G. (2009). Systèmes éducatifs et inégalités scolaires: une perspective internationale. *Sociologies*.
- Felouzis, G., Charmillot, S., & Fouquet-Chauprade, B. (2009). Systèmes éducatifs et inégalités scolaires en Suisse: une analyse de l'enquête PISA 2003.
- Felouzis, G., Charmillot, S., & Fouquet-Chauprade, B. (2011). Les inégalités scolaires en Suisse et leurs déclinaisons cantonales: l'apport de l'enquête Pisa 2003. *Swiss Journal of Sociology*, 37(1).
- Felouzis, G., & Hanhart, S. (Eds.). (2011). *Gouverner l'éducation par des nombres? Usages, débats et controverses*. Bruxelles/Paris: De Boeck.
- Felouzis, G., Maroy, C., & Van Zanten, A. (2013). *Les marchés scolaires : sociologie d'une politique publique d'éducation*. Paris: Presses Universitaires de France
- Field, A. (2009). *Discovering Statistics using SPSS (and sex and drugs and rock'n'roll)*. 3th edition. London: Sage.
- Figlio, D., & Loeb, S. (2011). School Accountability. . In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbooks in Economics* (Vol. 3, pp. 383-421). The Netherlands: North-Holland.
- Finnigan, K. S., & Gross, B. (2007). Do Accountability Policy Sanctions Influence Teacher Motivation? Lesson from Chicago's low performing schools. *American Educational Research Journal*, 44(3), 594 –629.
- Forum, W. E. (2015). *Incheon Declaration : Education 2030 : Towards inclusive and equitable quality education and lifelong learning for all*.
- Frederiksen, N. (1994). *The influence of minimum competency testing teaching and learning*. Retrieved from Princeton, NJ.:
- Frisbie, D. A. (1988). Reliability of scores from teacher- made tests. *Educational Measurement: Issues and Practice*, 7(1), 25-35.
- Fuentes, A. (2009). *Raising Education Outcomes in Spain*. Retrieved from
- Fuhrman, S. (1987). Education Policy: A new Context for Governance. *Publius: The Journal of Federalism*, 17(3), 131-143.
- Fuhrman, S. (1999). *The new accountability* Retrieved from Philadelphia, PA:
- Fuhrman, S., & Elmore, R. (2004). *Redesigning accountability systems for education*. New York: Teacher College Press.
- Gardner, D. P., Larsen, Y. W., Baker, W., Campbell, A., & Crosby, E. A. (1983). *À Nation At Risk: The Imperative For Educational Reform*: United States Department of Education.

- Glenn, C., & Galan, A. (2011). Educational Freedom and Accountability for School Quality in Spain. *Internal Journal of Education Law and Policy*, 7(1-2), 179-193.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Erlaun: Hillsdale, NJ.
- Grek, S. (2009). Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy*, 24(1), 23-37.
- Grodsky, E., Warren, J. R., & Kalogrides, D. (2009). State High School Exit Examinations and NAEP Long-Term Trends in Reading and Mathematics, 1971-2004. *Educational Policy*, 23(4), 589-614.
- Haccoun, R. R., & Cousineau, D. (2007). *Statistiques: Concepts et applications*. Montréal: Les Presses de l'Université de Montréal.
- Hair, J. F. J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis Seventh Edition*: Prentice Hall.
- Halinen, I. (2006). L'école de l'équité et de l'intégration. Le cas de la Finlande. *Revue internationale d'éducation de Sèvres*, 41, 75-86.
- Halinen, I., Kauppinen, J., & Yrjölä, P. (2006). Que savent les élèves en Finlande. *Revue internationale d'éducation de Sèvres*, 43, 105-104.
- Hallak, J. (1996). *Couts et dépenses en éducation. Principes de la planification*. Retrieved from Paris
- Hantrais, L. (1999). Contextualization in cross-national comparative research. *International Journal of Social Research Methodology*, 2(2), 93-108.
- Hanushek, E. A., & Raymond, M. E. (2005). Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management*, 24(2), 297-327.
- Harris, D. N., & Herrington, C. D. (2006). Accountability, Standards, and the Growing Achievement Gap: Lessons from the Past Half-Century. *American Journal of Education*, 112(2), 209-238.
- Ho, R. (2008). *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. Broken Sound Parkway: Taylor & Francis.
- Hood, C. (1991). A Public Management for All Seasons. *Public Administration*, 69, 3-19.
- Hood, C. (2007). Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles? *Public Money & Management*, 27(2), 95-102. doi:10.1111/j.1467-9302.2007.00564.x
- Houdayer, P., & Fontane, F. (2001). *Configuration et coordination des activités internationales de Recherche-Développement*. Paper presented at the Xième Conférence de l'Association Internationale de Management Stratégique.
- Howell, D. (2006). *Méthodes statistiques en sciences humaines. 6th Edition*. Bruxelles De Boeck.
- Hutmacher, W. (1993). Quand la réalité résiste à l'échec à la lutte contre l'échec scolaire. Analyse du redoublement dans l'enseignement primaire genevois. *Cahier du service de la recherche sociologique (SRS)*, 36.

- Hutmacher, W., Cochrane, D., & Bottani, N. (Eds.). (2001). *In the pursuit of equity in education. Using international indicators to compare equity policies*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Jaafar, S. B., & Anderson, S. (2007). Policy Trends and Tensions in Accountability for Educational Management and Services in Canada. *Alberta Journal of Educational Research*, 53(2), 207-227.
- Jacob, B. (2001). Getting tough ? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23(2), 99-121.
- Jacob, B. (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(761-796).
- Jencks, C. (1972). *L'inégalité. Influence de la famille et de l'école en Amérique*. Paris: Presses Universitaires de France.
- Kogan, M. (1988). *Education accountability : an analytic overview*. London ; Dover, NH, USA: Hutchinson Education.
- Kymlicka, W. (1999). *Les théories de la justice*. Paris: La Découverte.
- Ladd, H. F., & Lauen, D. L. (2010). Status versus Growth: The Distributional Effects of School Accountability. *Policies Journal of Policy Analysis and Management*, 29(3), 426-450.
- Lascoumes, P., & Le Galès, P. (2004). *Gouverner par les instruments*. Paris: SciencesPo.Les Presses.
- Lascoumes, P., & Simard, L. (2011). L'action publique au prisme de ses instruments. *Revue française de science politique*, 61(1), 5-22.
- Le Donné, N. (2017). L'éducation au prisme des enquêtes PISA. *Idées économiques et sociales*, 1(187), 17-26.
- Lee, J. (2007). Do national and state assessments converge for educational accountability ? A meta-analytic synthesis of multiple measures in Maine and Kentucky. *Applied Measurement in education*, 20(2), 1-33.
- Lee, J. (2008). Is test driven external accountability effective? Synthesizing the evidence from cross state Causal comparative and correlational studies. *Review of Educational Research*, 78(3), 608-644.
- Lee, J. (2010). Trick or treat : new ecology of education accountability system in the USA. *Journal of Education Policy*, 25(1), 73-93.
- Lee, J., & Wong, K. K. (2004). The impact of the accountability on racial and socioeconomic equity considering both school resources and achievement outcomes. *American Educational Research Journal*, 41(4), 797-832.
- Leithwood, K., & Earl, L. (2000). Educational Accountability Effects: An International Perspective. *Peabody Journal of Education*, 75(4), 1-18.
- Lessard, C. (2002). *L'État, le marché, le travail enseignant*, Montréal.
- Lessard, C., & Desjardins, P. D. (2011). Les politiques d'imputabilité en Ontario et au Québec: le point de vue des directions. *Education Comparée*, 5, 83-105.

- Lessard, C., & Meirieu, P. (2004). *L'obligation de résultats en éducation*. Québec: Presses de l'Université Laval.
- Lindle, J. C. (2009). Assessment Policy and Politics of information. In G. Sykes, B. Scheinder, D. N. Plank, & T. G. Ford (Eds.), *Handbook of education policy research* (pp. 319-329). New York & London: Routledge.
- Lingard, B., Martino, W., Rezai-Rashti, G., & Sellar, S. (2016). *Globalizing Educational Accountabilities*. New York: Routledge.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- MacCallum, R., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample Size in Factor Analysis. *Psychological Methods*, 4(1), 84-99.
- Marks, G. N. (2005). Cross-national differences and accounting for social class inequalities in education. *International sociology*, 20(4), 483-505.
- Maroy, C. (2005). Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe ? . *Cahiers de Recherche en Éducation et Formation*, 49.
- Maroy, C. (2006). *École, régulation et marché : une comparaison de six espaces scolaires locaux en Europe*. Paris: Presses Universitaires de France.
- Maroy, C. (2007). L'école à la lumière de la sociologie des organisations. *Cahiers de Recherche en Éducation et Formation*, 56.
- Maroy, C. (2008). Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe ? *Sociologie et sociétés*, 40(1), 31-55.
- Maroy, C. (2009). Enjeux, présupposés et implicites normatifs de la poursuite de l'efficacité dans les systèmes d'enseignement. In X. Dumay & V. Dupriez (Eds.), *L'efficacité dans l'enseignement : promesses et zones d'ombre* (1 ed., pp. 209-224). Bruxelles: De Boeck
- Maroy, C. (2017). Régulation In A. van Zanten (Ed.), *Dictionnaire de l'éducation*. Paris: Presses Universitaires de France.
- Maroy, C. (Ed.) (2013). *L'école à l'épreuve de la performance: Les politiques de régulation par les résultats : trajectoires nationales, usages locaux*: DeBoeck (à paraître).
- Maroy, C., Brassard, A., Mathou, C., Vaillancourt, S., & Voisin, A. (2015). *La mise en oeuvre de la politique de Gestion Axée sur les Résultats dans les commissions scolaires au Québec. Médiations et mécanismes d'institutionnalisation d'une nouvelle gestion de la pédagogie. Rapport de l'axe 2 du projet de recherche New-Age Québec (New Accountability and Governance in Education) & CRSH Savoir*. Montréal: Chaire de recherche du Canada sur les politiques éducatives.
- Maroy, C., Brassard, A., Mathou, C., Vaillancourt, S., & Voisin, A. (2017). La co-construction de la Gestion axée sur les résultats : les logiques de médiations des Commissions scolaires québécoises. *Mc Gill Journal of Education*, 52(1), 93-114.
- Maroy, C., Mathou, C., Vaillancourt, S., & Voisin, A. (2013). *"La Gestion axée sur les résultats" au Québec : récits d'action publique, intérêts des acteurs et médiations institutionnelles dans la fabrication d'une politique éducative*. Montréal: Chaire de recherche du Canada en politiques éducatives, Université de Montréal

- Maroy, C., Pons, X., & Dupuy, C. (2017). Vernacular Globalisations. Accountability in Education Policy in France and Quebec. *Journal of Education Policy*, 32(1), 100-122.
- Maroy, C., Vaillancourt, S., & Mathou, C. (2017). La gestion axée sur les résultats au coeur de l'école québécoise : l'autonomie professionnelle des enseignants sous pression. In Y. Dutercq & C. Maroy (Eds.), *Professionnalisme enseignant et politiques de responsabilisation* (pp. 35-55). Bruxelles: De Boeck.
- Maroy, C., & Voisin, A. (2013). Les transformations récentes des politiques d'accountability en éducation : enjeux et incidences des outils d'action publique. *Educação & Sociedade*, 124(34), 801-818.
- Maroy, C., & Voisin, A. (2014). Diversité des politiques d'accountability : l'incidence de l'instrumentation et des théories de la régulation des politiques. *Education Comparée*, 11, 31-57.
- Maroy, C., & Voisin, A. (2017a). Accountability. In A. R. Van Zanten, P. (Ed.), *Dictionnaire de l'éducation. 2^e Édition*. Paris: Presses Universitaires de France.
- Maroy, C., & Voisin, A. (2017b). *Monitoring and accountability in education – Opportunities, risks and innovations in the post-2015 period. Think piece on accountability policies for the Global Education Monitoring Report*. Paris: UNESCO.
- Maurice, M. (1989). Méthode comparative et analyse sociétale: Les implications théoriques des comparaisons internationales. *Sociologie du travail*, 31(2), 175-191.
- Mauss, M. (1968). *Essai de sociologie*. Paris: Éditions de minuit.
- McEwen, N. (1995). Introduction: Accountability in education in Canada. *Canadian Journal of Education/Revue canadienne de l'éducation*, 1-17.
- McNeil, L., Coppola, E., Radigan, J., & Vasquez Heilig, J. (2008). Avoidable Losses : High-Stake Accountability and the Dropout crisis. *Education Policy Analysis Archive*, 16(3).
- Meadwell, H. (2002). La théorie du choix rationnel et ses critiques. *Sociologie et sociétés*, 34(1), 117-124.
- MELS. (1997). *Prendre le virage du succès: l'école, tout un programme: énoncé de politique éducative*. Québec Ministère de l'éducation, du loisir et du sport.
- MEO. (2010). *Évaluation du rendement du personnel enseignant. Guide des exigences et des modalités*. Toronto: Ministère de l'éducation de l'Ontario.
- Meuret, D. (2012). Les effets de la régulation par les résultats (accountability) sur les politiques d'éducation aux États-Unis. *Education et sociétés*, 30(2), 75-87.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, 83(2), 340-363. doi:10.2307/2778293
- Mingat, A. (1988). Mesure et analyse de l'égalité et de l'équité en éducation. *Revue économique*, 39(1), 93-112.
- Mintrop, H., & Sunderman, G. (2009). Predictable failure of federal sanction-driven accountability for school improvement - And why we may retain it anyway. *Educational Researcher*, 38(5), 353-364.

- Mons, N. (2004). *De l'école unifiée aux écoles plurielles. Evaluation internationale des politiques de différenciation et de diversification de l'offre éducative.* (Doctorat en Sciences de l'Éducation), Université de Bourgogne.
- Mons, N. (2007). *Les nouvelles politiques éducatives. La France fait-elle les bons choix ?* Paris: Presses Universitaires de France.
- Mons, N. (2008). Élite scolaire, inégalités sociales et renouveau des filières dans l'école moyenne : une comparaison internationale. *Education et sociétés*, 1(21), 17-32.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie*, 169, 99-140.
- Mons, N. (2017). Efficacité des systèmes éducatifs. In A. van Zanten (Ed.), *Dictionnaire de l'éducation. 2^e Édition.* Paris: Presses Universitaires de France.
- Mons, N., & Dupriez, V. (2011). Les politiques d'accountability. *Recherche et formation*, 65.
- Mons, N., & Pons, X. (2006). *Les standards en éducation dans le monde francophone. Une analyse comparative.* Neuchâtel: Institut de recherche et de documentation pédagogique.
- Muller, P. (2000). L'analyse cognitive des politiques publiques: vers une sociologie politique de l'action publique. *Revue française de science politique*, 2, 189-208.
- Neave, G. (1988). On the cultivation of quality, efficiency and enterprise: an overview of recent trends in higher education in Western Europe, 1986-1988. *European Journal of Education*, 23(1-2), 7-23.
- Normand, R. (2006). L'école efficace ou l'horizon du monde comme laboratoire. *Revue française de pédagogie*, 154.
- OCDE. (2001). *Connaissances et compétences: des atouts pour la vie – Premiers résultats de PISA 2000.* Paris: OCDE.
- OCDE. (2007). *PISA 2006: science competencies for tomorrow's world, vol. 1.* Paris: OCDE.
- OCDE. (2009). *PISA Data Analysis Manual : SPSS and SAS, Second Edition.* Paris: OCDE.
- OCDE. (2010). *Les résultats du PISA 2009. Synthèse.* Paris: OCDE.
- OCDE. (2013a). *PISA 2012 Results : Excellence through equity. Giving every student the chance to succeed. Vol 2. .*
- OCDE. (2013b). *PISA 2012 Results : What makes schools successful ? Ressources, policies and practices. Vol IV.* Paris: OCDE.
- OCDE. (2014a). *PISA 2012 Technical Report.* Paris: OCDE.
- OCDE. (2014b). *Principaux résultats de l'Enquête PISA 2012. Ce que les élèves de 15 ans savent et ce qu'ils peuvent faire avec ce qu'ils savent.* Paris: OCDE.
- OCDE. (2014c). *Résultats du PISA 2012 : Savoirs et savoir-faire des élèves. Performance des élèves en mathématiques, en compréhension de l'écrit et en sciences. Volume 1.* Paris: OCDE.
- OCDE. (2016). *PISA 2015. Résultats à la loupe.* Paris: OCDE.
- Ozga, J. (2009). Governing education through data in England: From regulation to self-evaluation. *Journal of Education Policy*, 24(2), 149-162.

- Ozga, J., & Grek, S. (2010). Governing Education through Data: Scotland, England and the European Education Policy Space. *British Education Research Journal*, 36(6), 937-952.
- Ozga, J., & Grek, S. (2012). Governing through Learning. School Self-Evaluation as a Knowledge-based Regulatory Tool. *Recherches sociologiques et anthropologiques*, 43(2), 5-52.
- Parijs, V. (1991). *Qu'est-ce qu'une société juste ?*. Paris: Seuil.
- Paul, J.-J. (2005). Regard d'un économiste sur l'efficacité en éducation. In M. Demeuse, A. Baye, M. H. Straeten, & J. M. Nicaise, A. (Eds.), *Vers une école juste et efficace* (pp. 29-42). Bruxelles: De Boeck Supérieur.
- Pons, X. (2014). Le « pilotage par les résultats » en éducation en France : archéologie d'une formalisation administrative. *Education Comparée*, 11, 59-81.
- Pourtois, H. (2008). Pertinence et limites du principe d'égalité des chances en matière d'éducation scolaire. In V. Dupriez, J. F. Orienne, & M. Verhoven (Eds.), *De l'école au marché du travail, l'égalité des chances en question*. Berne: Peter Lang.
- Ranson, S. (2003). Public accountability in the age of neo-liberal governance. *Journal of Education Policy*, 18(5), 459-480. doi:10.1080/0268093032000124848
- Rapkin, B., & Luke, D. (1993). Cluster analysis in community research : epistemology and practice. *American Journal of Psychology*, 21(2), 247-277.
- Raveaud, M. (2007). L'éducation comparée: nouveaux débats pour des paradigmes bicentennaires. *Revue internationale de politique comparée*, 14(3), 377-384.
- Ravitch, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Book.
- Rawls, J. (2003). *La justice comme équité. Une reformulation de Théorie de la justice*, trad. B. Guillaume. Paris: La Découverte.
- Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics*, 92(5-6), 1394-1415.
- Rémond, M. (2006). Éclairages des évaluations internationales PIRLS et PISA sur les élèves français. *Revue française de pédagogie*, 157, 71-84.
- Rhoten, D., Carnoy, M., Chabran, M., & Elmore, R. (2003). The Conditions and Characteristics of Assessment and Accountability. The Case of Four States. In M. Carnoy, R. Elmore, & L. Siskin (Eds.), *The New Accountability. High Schools and High stakes testing* (pp. 13-54). New-York: Routledge.
- Rocher, T., & Le Donnée, N. (2012). Les aspirations professionnelles des élèves de 15 ans dans 57 pays: ambition et réalisme. *L'orientation scolaire et professionnelle*, 41(3).
- Rochex, J.-Y., & Kherroubi, M. (2004). La recherche en éducation et les ZEP en France (2). Apprentissages et exercice professionnel en ZEP : résultats, analyses, interprétations. *Revue française de pédagogie*, 146, 115-190.
- Rose, N. (1991). Governing by numbers : figuring out democracy. *Accounting Organizations and Society*, 16(7), 673-692.

- Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review*, 23(2), 143-151.
- Rowan, B. (2006). The school improvement industry in the United States: Why educational change is both pervasive and ineffectual. In H. D. Meyer, B. Rowan, & I. NetLibrary (Eds.), *The new institutionalism in education* (pp. 67-86). Albany: State University of New York Press.
- Rustique-Forrester, E. (2005). Accountability and the Pressures To Exclude: A Cautionary Tale from England. *Education Policy Analysis Archives*, 13(26).
- S.d. (2016). Classement PISA : pourquoi la France ne progresse-t-elle pas ? *Le monde* (7 décembre).
- Sahlberg, P. (2007). Education policies for raising student learning: the Finnish approach. *Journal of Education Policy*, 22(2), 147-171.
- Sahlberg, P. (2010). Rethinking accountability in a knowledge society. *Journal of Educational Change*, 11(45-61).
- Sahlberg, P. (2011). The Fourth Way of Finland. *Journal of Educational Change*, 12(2), 173-185.
- Sall, H. N., & De Ketele, J.-M. (1997). L'évaluation du rendement des systèmes éducatifs : apports des concepts d'efficacité, d'efficience et d'équité. *Mesure et Evaluation en Education*, 19(3), 119-142.
- Schnapper, D. (1999). *La compréhension sociologique: démarche de l'analyse typologique*. Paris: Presses universitaires de France.
- Schütz, G., West, M., & Wößmann, L. (2007). *School accountability, autonomie, choice and the equity of student achievement : international evidence from PISA 2003*. Retrieved from Paris
- Simola, H. (2005). The Finnish miracle of PISA : historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41(4), 455-470.
- Slavin, R., Karweit, N., & Madden, N. (1989). *Effective programs for students at risk*. Boston: Allyn & Bacon.
- Statistisches-Bundesamt. (2011). *Bildung und Kultur. Allgemeinbildende Schulen. Schuljahr 2009/2010*. Wiesbaden: Statistisches Bundesamt.
- Statistisches-Bundesamt. (2012). *Internationale Bildungsindikatoren im Ländervergleich*. Wiesbaden: Statistisches-Bundesamt.
- Suspitsyna, T. (2014). Cultural hierarchies in the discursive representations of China in the Chronicle of Higher Education. *Critical Studies in Education*, 56(1), 21-37. doi:10.1080/17508487.2015.971330
- Tabachnick, B. G., & Fidell, B. (2013). *Using multivariate Statistics*. USA Pearson Education.
- Thélot, C., & Vallet, L. A. (2000). La réduction des inégalités sociales devant l'école depuis le début du siècle. *Économie et statistique*, 334(1), 3-32.

- Théret, B. (1998). La régulation politique : le point de vue d'un économiste. In J. Commaille & B. Jobert (Eds.), *Les métamorphoses de la régulation politique*. Paris: Librairie de Droit et de Jurisprudence.
- Thoenig, J. C. (1998). L'usage analytique du concept de régulation. In J. Commaille & B. Jobert (Eds.), *Les métamorphoses de la régulation politique* (pp. 35-53). Paris: Librairie générale de droit et de jurisprudence.
- Väljärvi, J., Kupari, P., Linnakylä, P., Reinikainen, P., Sulkunen, S., Törnroos, J., & Arffman, I. (2002). *The Finnish success in Pisa-and some reasons behind it*. Retrieved from Jyväskylä:
- Van Daele, H. (1993). *L'Education comparée*. Collection « Que sais-je ? ». Paris: Presses Universitaires de France.
- Van Zanten, A. (2009a). *Choisir son école. Stratégies familiales et médiations locales* Paris Presses universitaires de France
- Van Zanten, A. (2009b). Le choix des autres. *Actes de la recherche en sciences sociales*, 5, 24-43.
- Vandenberghe, V. (2001). Nouvelles formes de régulation dans l'enseignement : origines, rôle de l'évaluation et enjeux en termes d'équité et d'efficacité. *Education et sociétés*, 8, 111-123.
- Verger, A., & Curran, M. (2014). New public management as a global education policy: its adoption and re-contextualization in a Southern European setting. *Critical Studies in Education*, 1-19. doi:10.1080/17508487.2014.913531
- Verger, A., & Parcerisa, L. (2017). *Monitoring and accountability in education – Opportunities, risks and innovations in the post-2015 period. Think piece on accountability policies for the Global Education Monitoring Report*. Paris: Unesco.
- Verhoeven, M., & Dupriez, V. (2008). Figures de l'égalité et politiques éducatives en Belgique francophone. In V. Dupriez, J. F. Orienne, & M. Verhoven (Eds.), *De l'école au marché du travail, l'égalité des chances en question*. Berne: Peter Lang.
- Vesely, A. (2013). Accountability in Central and Eastern Europe: concept and reality. *International Review of Administrative Sciences*, 79(2), 310-330.
- Vrignaud, P. (2006). La méthodologie dans PISA : la méthodologie est la réponse, mais quelle était la question ? *Revue française de pédagogie*(157), 27-41.
- Webb, P. T. (2005). The anatomy of accountability. *Journal of Education Policy*, 20(2), 189-208.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.
- Willms, J. D. (2004). *Variation in literacy skills among Canadian provinces: Findings from the OECD PISA*: Statistics Canada, Culture, Tourism and the Centre for Education Statistics.
- Wolf, I., & Janssens, F. J. G. (2007). Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, 33(3), 379-396.
- Wößmann, L. (2005). The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics*, 13(2), 143-169.

- Wößmann, L. (2007). International Evidence on School Competition, Autonomy, and Accountability: A Review. *Peabody Journal of Education*, 82(2), 473-494.
- Wößmann, L. (2011). Cross-country evidence on teacher performance pay. *Economics of Education Review*, 30, 404-418.
- Y., E., Pras, B., & Roux, E. (1993). *Études et recherches en marketing*. Paris: Nathan.
- Yerly, G. (2014). *Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignants. Analyse du regard des enseignants du primaire*. (Thèse de doctorat), Université de Fribourg.

ANNEXE 1

SYNTHÈSE DES ÉTUDES CORRÉLATIONNELLES ET CAUSALES PORTANT SUR LES EFFETS DES SABP SUR L'EFFICACITÉ ET L'ÉQUITÉ SCOLAIRE

(sources : Lee, 2008 & auteur sur base de la revue de littérature réalisée au chapitre 1 section 2)

	Contexte	<i>Accountability</i>	Efficacité	Équité	Données	Démarche	Effets ⁷²
Fredericks en (1994)	Comparaison inter-États (USA)	<i>Testing</i> basé sur les compétences minimums (MCT) Classification des États selon les enjeux liés aux MCT (enjeux forts, modérés, faibles)	Scores maths (agrégés) <i>Grades</i> 4, 8 et 12 Compétences de base et tâches complexes	non	NAEP 1978 et 1986	Comparatif gains de performances par âge dans États à forts et à faibles enjeux	Positifs pour <i>grade</i> 4 (tâches simples et complexes) Positif pour les <i>grade</i> 8 et 12 (tâches simples) Négatif <i>grade</i> 8th et 12 (tâches complexes) Taille d'effet — positive (nulle à faible) (Lee, 2008)
Jacob (2001)	Comparaison inter-États (USA)	Examens de sortie du secondaire (HSEEE) Classification des États présence vs absence	Scores maths et lecture <i>Grades</i> 8 et 12	Différentiel performance 1 ^{er} décile, 1 ^{er} quintile, moyenne distribution	NELS 1998 et 1992	Analyse des gains de performances par <i>grade</i> (année), risque de décrochage et liés aux HSEE (longitudinale)	Pas d'effet des HSEE après contrôle pour acquis antérieurs des élèves (8th) et caractéristiques des élèves, écoles, États Probabilité de décrochage (*25 %) : élèves 1 ^{er} quintile de la distribution des scores
Bishop <i>et al.</i> (2001)	Comparaison inter-États (USA)	Examens de fin de session MCT HSEEE SBR stratégies :	Scores maths et lecture <i>grade</i> 4 et 8 (NAEP) Scores maths, sciences,	Non	NAEP 1998 et 1996 NELS88	Analyse des gains de performances par <i>grade</i> (année) et risque de décrochage liés aux 1) examens de fin	Effets positifs des SBR Taille d'effet — positive (faible) Effets positifs de la combinaison MCT et Examens de fin de session Taille d'effet — positive (moyenne)

⁷² Les tailles d'effets sont ici abordées en termes génériques et basées sur les travaux de Lee (2008). Si le lecteur souhaite avoir une analyse plus fine des tailles d'effets, nous le renvoyons à ces travaux qui en proposent une analyse détaillée.

		publication des résultats des élèves, sanctions et récompenses pour les écoles	anglais, humanités, moyenne <i>grades</i> 8, 10, 12 (NELS88)			de session, 2) MCT ; 3) enjeux (longitudinale)	à faible) (Lee, 2008)
Amrein et Berliner (2002)	Comparaison inter-États (USA)	SABP à forts et à faibles enjeux SABP forts et faibles enjeux — indicateur composite (échelle de pression 2 à 8) : nombre d'instruments mobilisés (examen de graduation, publication des résultats, sanctions et récompenses pour les écoles)	Scores mathématiques et lecture (agrégés) Cohortes	Non	NAEP 1990 à 2000 et tests (ACT, SAT, AP)	Analyse des gains de performances par <i>grade</i> (année) et cohorte selon le degré de pression exercée par les États (comparatif un État versus autres États)	Effets inconstants/incohérents : positifs ou négatifs selon les cohortes (NAEP cohortes) Résultats sur les données du NAEP : effets constatés attribuables aux taux d'exclusion selon les auteurs Taille d'effets — mitigés (faibles à forts) (Lee, 2008)
Carnoy et Loeb (2002)	Comparaison inter-États (USA)	SABP forts et faibles enjeux — indice de pression (5 points) : instruments mis en place par les États (standards, testing et fréquence, HSEE, publication, sanctions,	Scores maths (agrégés) <i>4th</i> et <i>8th</i>	Différentiel perf. groupes ethniques : blancs, afro-américains et hispaniques	NAEP 1990, 1992, 1994, 1996, 2000	Analyse des gains de performances par <i>grade</i> (année) liés aux SABP [longitudinale]	Effets positifs liés à la force de l' <i>accountability</i> [échelle de pression] moyenne des scores, niveau le plus important : <i>8th</i> Gains pour tous les groupes d'élèves, plus importants pour les minorités <i>8th</i> Taille d'effet — positive [faible à forte] [Lee, 2008]
Hanushek et Raymond [2005]	Comparaison inter-États [USA]	SABP : sanctions, publication des résultats	Scores maths et lecture <i>4th</i> et <i>8th</i>	Différentiel perf. groupes ethnique : blancs, afro-américains et hispaniques	NAEP 1992-2000	Analyse des gains de performances par <i>grade</i> liés (année) aux SABP [longitudinale]	Effets positifs publication + sanction moyenne des scores Gains pour tous les groupes, mais plus importants pour blancs américains Augmentation de l'écart entre les groupes particulièrement blancs versus afro-Américains Taille d'effet — positive [faible à moyenne] [Lee, 2008]

Lee [2008]	Comparaison inter-États [USA]	SABP forts et faibles enjeux — indicateur composite [ACP] : politiques adoptées au moment de l'enquête Trois groupes « d'activistes » dans l' <i>accountability</i> : faible, modéré, fort	Scores maths [agrégés] 4 th et 8 th	Différentiel perf. groupes ethniques : blancs versus afro-américains et hispaniques	NAEP 1990, 1992, 1996, 2000	Analyse des gains de performances par <i>grade</i> (année) liés au SABP [longitudinale]	Effets positifs à court terme <i>grade</i> 8 : sanctions + publication Pas d'effets sur la réduction de l'écart de réussite après contrôle Taille d'effet — positive [faible] [Lee, 2008]
Lee [2009]	Comparaison inter-États [USA]	SABP : deux groupes voie de l'astuce ou de la conformité	Scores maths et lecture 4 th et 8 th [agrégés]	NA	NAEP vs tests à l'échelle des États	Analyse des gains de performances en fonction des voies choisies [longitudinale]	Différentiel de performance entre performances aux tests nationaux versus NAEP
Reback (2008)	Texas	MCT Notation et classements des écoles	Scores lecture, maths et « autres sujets » Grades 4, 8, 10	Différentiel perf. groupes ethniques : blancs, afro-américains et hispaniques	<i>Texas Assessment of Academic Skills</i> (TAAS) 1992-1998	Analyse des gains de performances [<i>tracking</i> élèves]	Maths : effets ++ plus importants pour les élèves dans groupe « pivot » [performance cruciale dans notation école] ; gains ++ pour les élèves dont les performances nécessitent une amélioration « moyenne » Lecture : gains pour les élèves « à la marge » Performances — aux matières « sans enjeux »
OCDE [2007]	Comparaison internationale	Instruments d' <i>accountability</i> : Information aux parents sur les performances des élèves, publication des résultats, usage des données de performance pour	Scores sciences	Non	PISA 2006	Comparaison gains de performances	Effets positifs, publication des résultats des élèves. Seule variable qui reste statistiquement significative après contrôle par facteurs démographiques et socioéconomiques

		évaluer les enseignants, les chefs d'établissements, etc.					
Schütz, West, Wößmann, [2007]	Comparaison internationale	Instruments d' <i>accountability</i> : 1) Examens de sortie, publication des résultats des élèves 2) Testing, usages des résultats des élèves pour promotion, suivi des progrès des écoles, comparer des performances entre écoles	Scores sciences et maths	Élèves 1 ^{er} décile distribution indicateur socioéconomique	PISA 2003	Comparaison gains de performances	Effets positifs des examens de sortie + publication + autonomie avec performances des élèves Pas d'effets d'interaction entre niveau socioéconomique et instruments d' <i>accountability</i> au niveau des établissements

ANNEXE 2

ÉCHANTILLON DE LA PRÉSENTE RECHERCHE

Systèmes éducatifs	Effectif total élèves	Effectifs pondérés
Autriche	4755	82225
Bulgarie	5282	54255
Suisse	11229	79621
République Tchèque	5327	82250
Danemark	7481	65642
Estonie	4779	11627
Finlande	8829	60047
France	4613	699779
Grèce	5125	96640
Croatie	5008	45506
Hongrie	4810	91179
Irlande	5016	54010
Islande	3508	4169
Italie	31073	521337
Lettonie	4306	16136
Lituanie	4618	33042
Luxembourg	5258	5521
Pays-Bas	4460	196262
Norvège	4686	59432
Pologne	4607	379545
Portugal	5722	96034
Roumanie	5074	140915
Slovaquie	4678	54636
Slovénie	5911	18418
Terreneuve	1312	4739
Prince_Edouard	1293	1596
Nouvelle Écosse	1374	10766
Nouveau-Brunswick	1784	7107
QC	4166	77924
On	3727	139054
Manitoba	2106	14307
Saskatchewan	1923	13039
Alberta	2103	36397
BC	1757	43059
CFB&Ger	3720	51179
Flandres	4877	66710
Angleterre, pays de Galles et Irlande du nord	9714	634338
Ecosse	2945	53898
Andalousie	1434	75553
Aragon	1393	9988
Asturias	1611	7125
Balearic Islands	1435	8385
Cantabria	1523	4334
Castille & Leon	1592	18422
Catalogne	1435	55833
Extramadura	1536	10399
Galice	1542	18287
La Rioja	1532	2566
Madrid	1542	48845

Murcia	1374	13115
Navarro	1530	5245
Pays Basque	4739	16143
Bade-Wurtemberg	727	110678
Bayern	784	111052
Hesse	354	51159
Basse-Saxe	469	73583
Rhénanie-nord	1209	178974
Rhénanie-Palatinat	252	37506
TOTAL (n=59)	222 351	4959533

ANNEXE 3

SYNTHÈSE DES DOMAINES D'ÉVALUATION DE L'ENQUÊTE PISA 2012

Synthèse des domaines d'évaluation de l'enquête PISA 2012

	CULTURE MATHÉMATIQUE	COMPRÉHENSION DE L'ÉCRIT	CULTURE SCIENTIFIQUE
Définitions	La culture mathématique est l'aptitude d'un individu à formuler, employer et interpréter les mathématiques dans un éventail de contextes, soit de se livrer à un raisonnement mathématique et d'utiliser des concepts, procédures, faits et outils mathématiques pour décrire, expliquer et prévoir des phénomènes. La culture mathématique aide les individus à comprendre le rôle que les mathématiques jouent dans le monde et à se comporter en citoyens constructifs, engagés et réfléchis, c'est-à-dire à poser des jugements et à prendre des décisions en toute connaissance de cause.	La compréhension de l'écrit est l'aptitude d'un individu à comprendre et utiliser des textes écrits, mais aussi à réfléchir à leur propos et à s'y engager. Cette capacité devrait permettre à chacun de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société.	La culture scientifique d'un individu renvoie à ses connaissances scientifiques et à sa capacité d'utiliser ces connaissances pour identifier les questions auxquelles la science peut apporter une réponse, pour acquérir de nouvelles connaissances, pour expliquer des phénomènes de manière scientifique et pour tirer des conclusions fondées sur des faits à propos de questions à caractère scientifique. Elle inclut la compréhension des traits caractéristiques de la science en tant que forme de recherche et de connaissances humaines, la conscience du rôle de la science et de la technologie dans notre environnement matériel, intellectuel et culturel, et la volonté de s'engager en qualité de citoyen réfléchi à propos de problèmes à caractère scientifique et touchant à des notions relatives à la science.
Contenus	Quatre idées majeures en rapport avec les nombres, l'algèbre et la géométrie : <ul style="list-style-type: none"> ▪ quantité ▪ espace et formes ▪ variations et relations ▪ incertitude et données 	Formats de l'écrit : <ul style="list-style-type: none"> ▪ textes continus – différents types de textes en prose (notamment de narration, d'information, d'argumentation, de description et d'instruction) ▪ textes non continus – présentant l'information sous d'autres formes (graphiques, diagrammes, formulaires et listes) 	Des connaissances et concepts scientifiques en rapport avec la physique, la chimie, la biologie et les sciences de la Terre et de l'espace doivent être appliqués au contenu des items et pas uniquement mémorisés.
Processus	<ul style="list-style-type: none"> ▪ formuler des situations de façon mathématique ▪ employer des concepts, faits, procédures et raisonnements mathématiques ▪ interpréter, appliquer et évaluer des résultats mathématiques (Les verbes « formuler », « employer » et « interpréter » désignent ces trois processus sous forme abrégée.)	<ul style="list-style-type: none"> ▪ localiser et extraire l'information ▪ comprendre le sens global d'un texte ▪ développer une interprétation ▪ réfléchir au contenu et à la forme d'un texte 	<ul style="list-style-type: none"> ▪ décrire, expliquer et prévoir des phénomènes scientifiques ▪ comprendre la démarche scientifique ▪ interpréter des données et des conclusions scientifiques
Contextes	Les situations dans lesquelles la culture mathématique est appliquée : <ul style="list-style-type: none"> ▪ contexte personnel ▪ contexte professionnel ▪ contexte sociétal ▪ contexte scientifique 	Les usages prévus des textes : <ul style="list-style-type: none"> ▪ usage personnel ▪ usage éducatif ▪ usage professionnel ▪ usage public 	Les situations dans lesquelles la culture scientifique est appliquée : <ul style="list-style-type: none"> ▪ contexte personnel ▪ contexte social ▪ contexte global Dans certaines applications de la science : <ul style="list-style-type: none"> ▪ la vie et la santé ▪ la Terre et l'environnement ▪ la technologie

ANNEXE 4

QUESTIONNAIRE « ÉCOLE » PISA 2012 (VERSION ANGLAIS)

(Seules les questions mobilisées dans nos analyses sont reprises dans l'annexe)

OECD Programme for International Student Assessment 2012

SCHOOL QUESTIONNAIRE FOR PISA 2012

Main Survey

No notes version

December 2011

Consortium:

Australian Council for Educational Research (ACER, Australia)

cApStAn Linguistic Quality Control (Belgium)

Deutsches Institut für Internationale Pädagogische Forschung (DIPF, Germany)

Educational Testing Service (ETS, USA)

Institutt for Lærerutdanning og Skoleutvikling (ILS, Norway)

Leibniz - Institute for Science and Mathematics Education (IPN, Germany)

National Institute for Educational Policy Research (NIER, Japan)

The Tao Initiative: CRP - Henri Tudor and Université de Luxembourg - EMACS
(Luxembourg)

Unité d'analyse des systèmes et des pratiques d'enseignement (aSPe, Belgium)

Westat (USA)

SECTION D: SCHOOL INSTRUCTION CURRICULUM AND ASSESSMENT

SC18

Q In your school, are assessments of students in <national modal grade for 15-year-olds> used for any of the following purposes?

(Please tick only one box in each row.)

- | | <i>Yes</i> | <i>No</i> |
|--|---------------------------------------|---------------------------------------|
| a) To inform parents about their child's progress | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| b) To make decisions about students' retention or promotion | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| c) To group students for instructional purposes | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| d) To compare the school to <district or national> performance | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| e) To monitor the school's progress from year to year | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| f) To make judgements about teachers' effectiveness | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| g) To identify aspects of instruction or the curriculum that could be improved | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| h) To compare the school with other schools | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |

SC19

Q In your school, are achievement data used in any of the following <accountability procedures>?

Achievement data include aggregated school or grade-level test scores or grades, or graduation rates.

(Please tick one box in each row.)

- | | <i>Yes</i> | <i>No</i> |
|--|---------------------------------------|---------------------------------------|
| a) Achievement data are posted publicly (e.g. in the media) | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |
| b) Achievement data are tracked over time by an administrative authority | <input type="checkbox"/> ₁ | <input type="checkbox"/> ₂ |

Q To what extent have appraisals of and/or feedback to teachers directly led to the following?

(Please tick one box in each row.)

	<i>No change</i>	<i>A small change</i>	<i>A moderate change</i>	<i>A large change</i>
a) A change in salary	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
b) A financial bonus or another kind of monetary reward	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
c) Opportunities for professional development activities	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
d) A change in the likelihood of career advancement	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
e) Public recognition from you	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
f) Changes in work responsibilities that make the job more attractive	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄
g) A role in school development initiatives (e.g. curriculum development group, development of school objectives)	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂	<input type="checkbox"/> ₃	<input type="checkbox"/> ₄

Q Which of the following measures aimed at quality assurance and improvement do you have in your school?

(Please tick one box in each row.)

	<i>Yes</i>	<i>No</i>
a) Written specification of the school's curricular profile and educational goals	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
b) Written specification of student performance standards	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
c) Systematic recording of data including teacher and student attendance and graduation rates, test results and professional development of teachers	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
d) Internal evaluation/self-evaluation	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
e) External evaluation	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
f) Seeking written feed-back from students (e.g. regarding lessons, teachers or resources)	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
g) Teacher mentoring	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
h) Regular consultation aimed at school improvement with one or more experts over a period of at least six months	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂
i) Implementation of a standardised policy for mathematics (i.e. school curriculum with shared instructional materials accompanied by staff development and training)	<input type="checkbox"/> ₁	<input type="checkbox"/> ₂

ANNEXE 5

STATISTIQUES DESCRIPTIVES POUR LES VARIABLES D'INTÉRÊT DE LA BASE DE DONNÉES « ÉCOLE » DU PISA 2012 – proportions d'élèves (de 15 ans) scolarisés dans des écoles dont le chef d'établissement a déclaré divers instruments de régulation par les résultats en usage au sein de son établissement scolaire, comparatif des systèmes éducatifs de l'échantillon

	Dispositifs d'assurance qualité & amélioration écoles (SC39)					Suivi des performances des écoles & publication des résultats (SC19)	
	Experts externes (Q09)	Évaluation Externe (Q06)	Évaluation interne (Q05)	Enregistrement données (Q03)	Spécification standards (Q02)	Suivi autorité administrative (Q02)	Publication (Q01)
Alberta	77,0	78,8	90,7	91,9	91,5	97,47	69,2
Andalousie	28,6	82,6	94,5	92,3	80,4	80,51	20,3
Angleterre, Pays de Galles & Irlande du nord	74,2	83,6	92,1	92,0	85,0	83,02	80,0
Aragon	18,1	58,3	66,9	89,9	70,7	80,47	6,4
Asturies	25,8	76,4	91,6	96,1	83,5	76,96	7,3
Autriche	52,6	19,5	83,4	72,4	53,9	57,47	5,54
Bade-Wurtemberg	23,5	45,4	74,3	56,3	53,1	6,82	0,0
Basse-Saxe	13,2	77,6	77,0	52,7	73,5	21,69	12,6
Bayern	10,4	63,9	47,0	68,9	43,6	38,44	3,4
Belgique CFB&GER	22,0	71,1	59,6	51,4	38,7	32,1	4,7
Belgique Flandres	50,5	62,4	89,2	91,4	52,4	62,96	1,7
Bulgarie	69,0	94,7	97,3	98,0	78,2	87,63	55,1
Cantabrie	16,2	81,1	76,8	93,0	80,5	85,85	20,2

Castille et Léon	18,6	55,2	78,6	96,6	76,8	81,28	12,1
Catalogne	35,8	92,1	82,1	88,3	70,9	92,55	10,7
Colombie-Britannique	47,2	45,8	69,5	85,2	84,9	90,85	56,9
Croatie	57,1	81,3	91,6	95,1	68,1	87,49	25,3
Danemark	41,6	49,6	74,7	68,4	32,3	61,23	34,8
Ecosse	65,3	89,5	97,1	94,1	79,0	84,59	79,9
Estonie	37,2	76,7	99,0	95,0	87,9	78,2	34,8
Extremadura	12,6	72,4	69,1	94,7	71,3	65,46	16,7
Finlande	9,8	50,8	94,6	73,2	74,1	47,16	1,6
France	18,8	47,1	55,2	67,9	22,7	70,44	42,7
Galice	14,6	61,8	60,1	70,2	71,2	57,31	7,1
Grèce	76,7	5,6	32,1	67,9	37,2	56,83	26,9
Hesse	10,5	51,1	75,3	62,7	41,5	25,6	27,0
Hongrie	16,9	55,9	94,3	77,8	88,2	56,7	47,2
Iles Baléares	28,3	61,4	79,8	73,6	59,8	71,9	6,2
Irlande	43,4	73,3	74,1	78,9	43,3	43,6	18,3
Islande	39,9	75,6	95,8	91,6	81,2	76,2	30,7
Italie	21,3	31,3	71,1	48,3	79,2	27,7	37,7
La Rioja	32,1	82,8	72,6	93,0	75,1	63,1	16,2
Lettonie	22,7	81,5	96,7	96,5	84,2	57,0	32,1
Lituanie	40,0	55,2	94,4	97,4	76,7	75,1	31,5
Luxembourg	36,7	35,6	64,9	67,4	41,6	67,3	14,0
Madrid	20,3	83,2	78,3	91,2	94,5	90,8	17,9
Manitoba	51,2	63,3	88,3	82,8	61,4	75,6	10,0
Murcia	20,8	48,1	79,4	84,3	77,4	62,5	12,8
Navarre	35,0	89,4	81,2	88,8	60,3	79,0	24,9

Norvège	28,8	48,4	56,3	77,2	67,3	79,9	50,9
Nouveau Brunswick	64,4	87,4	75,6	80,3	87,4	96,1	61,7
Nouvelle Ecosse	88,6	80,5	96,2	88,7	73,5	87,1	50,2
Ontario	88,0	72,7	87,9	93,2	89,0	94,4	72,7
Pays Basque	53,7	82,8	76,9	94,5	76,3	86,4	13,7
Pays-Bas	41,2	71,1	79,2	86,8	74,5	72,4	79,8
Pologne	38,5	75,6	96,3	98,2	79,2	75,4	46,3
Portugal	28,0	82,9	94,6	93,5	71,7	85,4	50,5
Prince Édouard	80,7	68,9	85,4	60,4	73,6	35,3	10,4
Québec	38,2	33,4	61,1	83,7	77,5	88,7	51,9
R-Tchèque	22,2	51,4	78,7	68,7	62,8	47,6	36,8
Rhénanie-nord	13,3	33,2	47,0	67,1	71,5	55,0	12,8
Rhénanie-Palatinat	17,5	53,4	64,0	61,9	69,1	0,0	0,0
Roumanie	66,2	83,6	87,7	88,6	86,6	69,8	67,9
Saskatchewan	68,3	50,1	83,4	83,2	70,5	93,0	31,2
Slovaquie	53,1	36,8	93,9	92,8	78,8	80,3	77,0
Slovénie	37,3	30,5	86,7	81,3	89,5	59,8	49,9
Suède	31,5	64,2	89,1	93,7	93,7	91,4	79,4
Suisse	25,0	58,0	77,8	58,2	39,6	48,8	5,4
Terre-Neuve et Labrador	66,8	94,4	98,4	94,6	84,7	96,5	52,5

	Usages des résultats des élèves par les établissements scolaires (SC18)			
	Comparaison performances autres écoles (Q08)	Évaluation performance enseignants (Q06)	<i>Monitoring</i> performance (Q05)	Comparaison performances nationales (Q0)4
Alberta	57,4	45,0	89,0	75,2
Andalousie	26,3	52,9	94,5	36,0
Angleterre, Pays de Galles & Irlande du nord	81,5	81,4	91,9	88,2
Aragon	14,9	27,3	72,1	23,4
Asturies	18,3	45,4	87,9	27,3
Autriche	25,7	38,4	61,4	27,9
Bade-Wurtemberg	15,6	6,0	33,1	30,5
Basse-Saxe	14,8	8,0	45,8	29,3
Bayern	25,0	37,6	52,7	48,2
Belgique CFB&GER	10,4	23,1	48,8	17,5
Belgique Flandres	22,8	41,2	63,2	25,9
Bulgarie	83,2	90,7	92,4	83,8
Cantabrie	39,5	59,5	91,8	56,3
Castille et Leon	28,5	59,1	93,1	30,8
Catalogne	43,9	51,4	90,2	64,0
Colombie-Britannique	66,6	29,1	89,0	91,7
Croatie	62,2	55,4	94,6	65,2
Danemark	46,6	23,8	49,7	48,0

Ecosse	84,7	73,6	96,5	89,9
Estonie	54,4	65,1	77,6	64,4
Extremadura	22,0	39,6	90,8	25,5
Finlande	20,0	15,3	58,8	45,2
France	37,6	21,0	68,5	58,0
Galice	23,8	43,9	75,3	25,0
Grèce	21,4	13,7	55,2	16,8
Hesse	26,5	16,3	62,5	34,5
Hongrie	69,0	56,3	90,2	76,5
Iles Baléares	46,5	54,2	87,1	56,0
Irlande	28,7	41,7	78,0	69,4
Islande	63,3	37,6	85,8	74,1
Italie	34,3	27,4	76,8	60,8
La Rioja	34,1	51,8	76,6	34,2
Lettonie	83,9	91,3	98,5	91,2
Lituanie	59,3	73,3	93,7	60,7
Luxembourg	39,8	22,3	72,3	74,3
Madrid	83,5	61,1	98,3	79,7
Manitoba	23,1	26,9	70,8	45,7
Murcia	7,3	47,6	85,2	18,3
Navarre	63,5	44,4	92,0	72,9
Norvège	45,4	28,3	78,6	64,0
Nouveau Brunswick	48,9	40,3	86,0	69,0
Nouvelle Ecosse	35,8	30,8	88,7	73,8
Ontario	64,1	21,2	96,4	87,8
Pays Basque	40,5	55,1	81,2	44,1

Pays-Bas	55,7	59,1	77,2	59,8
Pologne	58,2	77,3	95,1	57,0
Portugal	60,4	48,6	92,3	81,8
Prince Édouard	12,8	39,7	54,6	35,0
Québec	70,3	37,2	92,7	78,9
R-Tchèque	52,8	52,3	71,1	48,4
Rhénanie-nord	26,6	27,5	59,4	40,1
Rhénanie-Palatinat	16,0	7,5	51,9	16,9
Roumanie	69,1	74,8	72,4	67,6
Saskatchewan	53,2	25,0	85,8	84,6
Slovaquie	68,7	68,0	69,5	63,3
Slovénie	42,3	36,0	86,3	55,4
Suède	84,0	43,0	95,6	89,3
Suisse	24,2	33,6	44,3	37,9
Terre-Neuve et Labrador	79,4	53,4	94,9	88,6

	Incitatifs pour les enseignants (suites et conséquences articulés aux évaluations du corps enseignant) (SC31)					
	Attribution responsabilités professionnels (Q06)	Reconnaissance publique chef d'établissement (Q05)	Avancement de carrière (Q04)	Développement professionnel (Q03)	Bonus financier (Q02)	Changement de salaire (Q01)
Alberta	53,9	76,4	61,1	67,3	7,3	10,6
Andalousie	54,2	57,5	13,8	52,9	10,3	10,3
Angleterre, Pays de Galles & Irlande du nord	75,1	79,6	80,2	88,8	15,4	64,7
Aragon	34,2	47,4	15,6	41,4	0,0	0,0
Asturies	57,7	75,1	25,2	68,4	11,2	7,1
Autriche	43,3	73,3	29,7	34,7	8,0	3,3
Bade-Wurtemberg	35,8	25,9	36,6	44,3	2,8	0,0
Basse-Saxe	26,4	34,2	12,1	34,1	0,0	4,0
Bayern	42,9	54,0	65,0	61,7	32,5	21,8
Belgique CFB&GER	39,3	58,8	6,3	56,3	0,0	0,0
Belgique Flandres	53,1	63,1	33,3	70,4	1,7	0,6
Bulgarie	81,2	93,3	84,8	89,6	85,2	29,3
Cantabrie	49,2	60,7	20,2	57,7	7,5	6,2
Castille et Léon	49,2	71,7	14,1	41,4	5,2	11,0
Catalogne	58,0	83,2	45,1	42,0	11,6	10,0
Colombie-Britannique	43,2	62,9	33,1	69,9	1,2	1,2
Croatie	80,4	97,3	90,1	87,2	26,5	15,3
Danemark	47,5	66,8	12,4	56,8	6,2	3,6

Ecosse	64,5	71,9	73,2	95,0	0,2	1,0
Estonie	69,4	91,8	57,3	78,7	69,7	37,8
Extremadura	43,8	54,8	21,6	44,3	7,3	3,3
Finlande	66,4	74,2	26,3	69,6	22,5	18,8
France	54,0	71,4	57,4	56,9	18,2	38,2
Galice	26,2	36,8	11,0	15,4	6,3	9,8
Grèce	49,7	68,2	39,1	48,3	22,1	22,6
Hesse	45,8	51,9	33,2	77,5	0,0	0,0
Hongrie	84,5	95,9	72,7	66,1	80,4	21,6
Iles Baléares	54,6	66,6	40,6	46,1	9,9	12,1
Irlande	36,1	62,9	24,8	47,7	1,1	1,2
Islande	76,6	71,3	27,6	77,9	16,7	17,8
Italie	73,7	56,7	30,9	60,6	34,3	14,5
La Rioja	40,3	53,7	11,2	27,0	2,8	0,0
Lettonie	77,2	91,6	63,1	85,0	34,0	42,8
Lituanie	63,4	95,4	62,9	87,5	47,1	44,9
Luxembourg	52,3	73,6	17,3	44,1	2,1	2,1
Madrid	61,4	74,5	25,7	47,8	14,1	11,6
Manitoba	43,1	67,1	29,3	75,6	1,6	1,6
Murcia	54,7	66,3	24,9	42,6	3,9	3,9
Navarre	51,8	66,2	31,9	47,0	6,0	11,1
Norvège	70,3	71,3	45,3	76,0	2,9	7,7
Nouveau Brunswick	53,3	81,6	59,3	76,2	2,6	7,6
Nouvelle Ecosse	59,2	72,7	52,0	80,8	0,0	0,0
Ontario	34,4	75,9	45,8	82,6	0,1	0,7
Pays Basque	62,5	68,0	30,9	53,3	15,0	10,5

Pays-Bas	63,0	78,1	59,4	77,1	22,6	18,7
Pologne	59,5	90,8	55,6	72,9	80,7	33,6
Portugal	60,6	55,7	39,9	43,8	10,7	20,5
Prince Édouard	45,9	71,3	30,8	75,8	0,2	0,2
Québec	46,6	65,2	32,7	72,0	7,7	5,3
R-Tchèque	50,4	75,9	47,5	67,7	68,9	59,0
Rhénanie-nord	52,2	48,4	47,5	43,5	1,8	5,2
Rhénanie-Palatinat	14,8	37,7	46,1	53,6	7,3	0,0
Roumanie	72,8	76,2	71,6	66,0	33,1	29,8
Saskatchewan	59,2	64,1	51,6	87,8	1,1	1,6
Slovaquie	80,3	94,3	71,2	84,2	83,0	49,0
Slovénie	86,0	90,4	80,5	80,9	50,4	40,4
Suède	80,5	87,2	59,1	90,2	18,4	84,8
Suisse	35,4	39,5	19,5	52,6	15,8	10,9
Terre-Neuve et Labrador	48,0	66,9	43,4	89,4	0,0	0,0

ANNEXE 6

DESCRIPTIF DES TYPES DE SYSTÈME D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE PAR COMPOSANTE INSTRUMENTALE ET RÉSULTATS DES ANALYSES DE VARIANCE À UN FACTEUR DE CLASSIFICATION (ANOVA)

Caractéristiques descriptives des types empiriques de SABP, moyennes et écart-type pour chaque composante instrumentale									
		N (systèmes éducatifs)	Moyenne	Ecart type	Erreur standard	Intervalle de confiance à 95 % pour la moyenne		Minimum	Maximum
						Borne inférieure	Borne supérieure		
Évaluation performance	Type 1	20	-1,09	,65	,15	-1,40	-,79	-2,62	-,06
	Type 2	17	,81	,55	,13	,53	1,09	-,57	1,63
	Type 3	8	,49	,49	,17	,08	,90	-,25	1,28
	Type 4	14	,30	,63	,17	-,07	,67	-,89	1,18
	Total	59	,00	1,00	,13	-,26	,26	-2,62	1,63
Incitatifs	Type 1	20	-,36	,50	,11	-,59	-,13	-,96	,77
	Type 2	17	-,39	,41	,10	-,60	-,17	-1,12	,62
	Type 3	8	-,96	,35	,12	-1,25	-,67	-1,57	-,45
	Type 4	14	1,53	,57	,15	1,20	1,86	,43	2,40
	Total	59	,00	1,00	,13	-,26	,26	-1,57	2,40
Comparaison externe & amélioration interne	Type 1	20	-,13	,67	,15	-,44	,19	-1,28	1,10
	Type 2	17	-,89	,56	,14	-1,18	-,61	-1,80	,23
	Type 3	8	1,67	,32	,11	1,40	1,93	1,39	2,32
	Type 4	14	,31	,69	,19	-,09	,71	-,68	1,55
	Total	59	,00	1,00	,13	-,26	,26	-1,80	2,32

Test d'homogénéité des variances (Levene)				
	Statistique de Levene	ddl1	ddl2	Sig.
Évaluation performance	,383	3	55	,766
Incitatifs	,824	3	55	,486
Comparaison externe & amélioration interne	2,823	3	55	,047

ANOVA						
		Somme des carrés	ddl	Carré moyen	F	Sig.
Évaluation performance	Inter-groupes	38,137	3	12,712	35,201	,000
	Intragroupes	19,863	55	,361		
	Total	58,000	58			
Incitatifs	Inter-groupes	45,457	3	15,152	66,442	,000
	Intragroupes	12,543	55	,228		
	Total	58,000	58			
Comparaison externe & amélioration interne	Inter-groupes	37,553	3	12,518	33,670	,000
	Intragroupes	20,447	55	,372		
	Total	58,000	58			

Résultats des tests de comparaisons multiples (Tukey et Bonferroni)								
Variable dépendante		(I) Types de SABP	(J) Types de SABP	Différence moyenne (I-J)	Erreur standard	Sig.	Intervalle de confiance à 95 %	
							Borne inférieure	Borne supérieure
Évaluation performance	Tukey	1	2	-1,90142554 [*]	,19824404	,000	-2,4266441	-1,3762070
			3	-1,57933666 [*]	,25139566	,000	-2,2453726	-,9133007
			4	-1,39081581 [*]	,20941085	,000	-1,9456192	-,8360124
		2	1	1,90142554 [*]	,19824404	,000	1,3762070	2,4266441
			3	,32208888	,25765543	,598	-,3605315	1,0047092
			4	,51060973	,21688580	,098	-,0639974	1,0852169
		3	1	1,57933666 [*]	,25139566	,000	,9133007	2,2453726
			2	-,32208888	,25765543	,598	-1,0047092	,3605315
			4	,18852085	,26634287	,894	-,5171156	,8941573
		4	1	1,39081581 [*]	,20941085	,000	,8360124	1,9456192
			2	-,51060973	,21688580	,098	-1,0852169	,0639974
			3	-,18852085	,26634287	,894	-,8941573	,5171156
	Bonferroni	1	2	-1,90142554 [*]	,19824404	,000	-2,4440283	-1,3588228
			3	-1,57933666 [*]	,25139566	,000	-2,2674177	-,8912556
			4	-1,39081581 [*]	,20941085	,000	-1,9639826	-,8176490
		2	1	1,90142554 [*]	,19824404	,000	1,3588228	2,4440283
			3	,32208888	,25765543	1,000	-,3831255	1,0273032
			4	,51060973	,21688580	,133	-,0830163	1,1042358
		3	1	1,57933666 [*]	,25139566	,000	,8912556	2,2674177
			2	-,32208888	,25765543	1,000	-1,0273032	,3831255
			4	,18852085	,26634287	1,000	-,5404714	,9175131
		4	1	1,39081581 [*]	,20941085	,000	,8176490	1,9639826
			2	-,51060973	,21688580	,133	-1,1042358	,0830163
			3	-,18852085	,26634287	1,000	-,9175131	,5404714
Incitatifs	Tukey	1	2	,02767191	,15753615	,998	-,3896971	,4450409
			3	,60422350 [*]	,19977349	,019	,0749529	1,1334941
			4	-1,89232639 [*]	,16640995	,000	-2,3332052	-1,4514476
		2	1	-,02767191	,15753615	,998	-,4450409	,3896971
			3	,57655158 [*]	,20474787	,033	,0341021	1,1190011
			4	-1,91999831 [*]	,17234997	,000	-2,3766143	-1,4633823
		3	1	-,60422350 [*]	,19977349	,019	-1,1334941	-,0749529
			2	-,57655158 [*]	,20474787	,033	-1,1190011	-,0341021
			4	-2,49654989 [*]	,21165141	,000	-3,0572893	-1,9358105
		4	1	1,89232639 [*]	,16640995	,000	1,4514476	2,3332052
			2	1,91999831 [*]	,17234997	,000	1,4633823	2,3766143
			3	2,49654989 [*]	,21165141	,000	1,9358105	3,0572893
	Bonferroni	1	2	,02767191	,15753615	1,000	-,4035115	,4588554
			3	,60422350 [*]	,19977349	,023	,0574346	1,1510124
			4	-1,89232639 [*]	,16640995	,000	-2,3477978	-1,4368550
		2	1	-,02767191	,15753615	1,000	-,4588554	,4035115
			3	,57655158 [*]	,20474787	,040	,0161476	1,1369556
			4	-1,91999831 [*]	,17234997	,000	-2,3917278	-1,4482688
		3	1	-,60422350 [*]	,19977349	,023	-1,1510124	-,0574346
			2	-,57655158 [*]	,20474787	,040	-1,1369556	-,0161476
			4	-2,49654989 [*]	,21165141	,000	-3,0758492	-1,9172506
		4	1	1,89232639 [*]	,16640995	,000	1,4368550	2,3477978
			2	1,91999831 [*]	,17234997	,000	1,4482688	2,3917278

Comparaison externe & amélioration interne	Tukey	1	3	2,49654989 [*]	,21165141	,000	1,9172506	3,0758492	
			2	,76838370 [*]	,20114002	,002	,2354927	1,3012747	
			3	-1,79400709 [*]	,25506807	,000	-2,4697726	-1,1182416	
			4	-,43978826	,21246995	,176	-1,0026963	,1231198	
		2	1	-,76838370 [*]	,20114002	,002	-1,3012747	-,2354927	
			3	-2,56239079 [*]	,26141929	,000	-3,2549829	-1,8697987	
			4	-1,20817196 [*]	,22005409	,000	-1,7911731	-,6251709	
		3	1	1,79400709 [*]	,25506807	,000	1,1182416	2,4697726	
			2	2,56239079 [*]	,26141929	,000	1,8697987	3,2549829	
			4	1,35421883 [*]	,27023364	,000	,6382744	2,0701633	
		4	1	,43978826	,21246995	,176	-,1231198	1,0026963	
			2	1,20817196 [*]	,22005409	,000	,6251709	1,7911731	
			3	-1,35421883 [*]	,27023364	,000	-2,0701633	-,6382744	
		Bonferroni	1	2	,76838370 [*]	,20114002	,002	,2178545	1,3189128
				3	-1,79400709 [*]	,25506807	,000	-2,4921397	-1,0958745
				4	-,43978826	,21246995	,259	-1,0213280	,1417514
	2		1	-,76838370 [*]	,20114002	,002	-1,3189128	-,2178545	
			3	-2,56239079 [*]	,26141929	,000	-3,2779070	-1,8468746	
			4	-1,20817196 [*]	,22005409	,000	-1,8104698	-,6058742	
	3		1	1,79400709 [*]	,25506807	,000	1,0958745	2,4921397	
			2	2,56239079 [*]	,26141929	,000	1,8468746	3,2779070	
			4	1,35421883 [*]	,27023364	,000	,6145774	2,0938603	
	4		1	,43978826	,21246995	,259	-,1417514	1,0213280	
			2	1,20817196 [*]	,22005409	,000	,6058742	1,8104698	
			3	-1,35421883 [*]	,27023364	,000	-2,0938603	-,6145774	
	*. La différence moyenne est significative au niveau 0.05.								

ANNEXE 7

CARACTÉRISTIQUES DE L'ENSEMBLE DE L'ÉCHANTILLON - indicateurs d'efficacité et d'équité scolaire (discipline de référence les mathématiques), composantes instrumentales des SABP et variables de contrôle

	Minimum	Maximum	Moyenne	Écart- type	Asymétrie	Kurtose
Efficacité	439	537	497	21,22	-0,49	0,44
Inégalités de résultats	1,50	1,77	1,62	0,06	0,33	-,39
Inégalités des chances de résultats	0,06	0,25	0,14	0,04	0,25	-,44
Inégalités des acquis de base	10,54	43,76	20,75	6,95	1,17	1,76
Composante (logique) instrumentale d'évaluation selon la performance	-2,62	1,63	,00	1,00	-0,51	-0,61
Composante (logique) d'incitatifs pour les enseignants	-1,57	2,40	,00	1,00	0,93	-0,23
Composante (logique) de comparaison des résultats et d'amélioration des performances des établissements scolaires	-1,80	2,32	,00	1,00	0,36	-0,74
Dépenses publiques en éducation	3,10	8,80	5,21	1,27	0,96	0,37
Variation du niveau socioéconomique des élèves	0,56	1,41	0,84	,17	0,74	0,64
Stratification horizontale & verticale	-1,02	1,89	,00	0,70	0,27	-0,75

ANNEXE 8

DESCRIPTIF DES DEGRÉS D'EFFICACITÉ ET D'ÉQUITÉ SCOLAIRE PAR TYPE DE SYSTÈME D'ACCOUNTABILITY BASÉS SUR LA PERFORMANCE ET RÉSULTATS DES ANALYSES DE VARIANCE À UN FACTEUR DE CLASSIFICATION (ANOVA)

TYPES DE SABP	Efficacité	Inégalités de résultats	Inégalités des chances de résultats	Inégalités des acquis de base	
1 (n = 20)	<i>Accountability</i> de faible intensité et peu instrumentée				
	Moyenne	504,75	1,62	,15	19,13
	Ecart type	21,16	,05	,04	5,76
	Asymétrie	-,37	,02	-,19	1,16
	Erreur standard d'asymétrie	,51	,51	,51	,51
	Kurtosis	,43	-1,17	-,53	2,22
	Erreur standard de Kurtosis	,99	,99	,99	,99
	Minimum	452,97	1,54	,07	11,17
	Maximum	536,60	1,70	,22	35,69
2 (n = 17)	<i>Accountability</i> axée sur l'évaluation de la performance				
	Moyenne	492,70	1,62	,14	21,29
	Ecart type	17,06	,06	,03	5,85
	Asymétrie	-,47	,16	-,13	,43
	Erreur standard d'asymétrie	,55	,55	,55	,55
	Kurtosis	-,42	-1,08	-,77	-,32
	Erreur standard de Kurtosis	1,06	1,06	1,06	1,06
	Minimum	461,37	1,54	,08	12,27
	Maximum	518,75	1,71	,20	32,96
3 (n = 8)	<i>Accountability</i> axée sur l'évaluation de la performance et la comparaison des résultats des établissements scolaires				
	Moyenne	504,95	1,56	,10	16,71
	Ecart type	12,35	,02	,04	3,21
	Asymétrie	,49	,38	1,43	,05
	Erreur standard d'asymétrie	,75	,75	,75	,75
	Kurtosis	-1,62	-,07	1,84	-1,36
	Erreur standard de Kurtosis	1,48	1,48	1,48	1,48
	Minimum	490,41	1,53	,06	12,32
	Maximum	521,99	1,61	,18	21,31
4 (n=14)	<i>Accountability</i> de forte intensité prévoyant des incitatifs directs pour les enseignants				
	Moyenne	486,82	1,64	,16	24,71
	Ecart type	25,42	,07	,05	9,39
	Asymétrie	-,40	,03	,52	,66
	Erreur standard d'asymétrie	,60	,60	,60	,60
	Kurtosis	-,16	,50	-,72	,30
	Erreur standard de Kurtosis	1,15	1,15	1,15	1,15
	Minimum	438,74	1,50	,09	10,54
	Maximum	522,97	1,77	,25	43,76

Test d'homogénéité des variances (Levene)				
	Statistique de Levene	ddl1	ddl2	Sig.
Efficacité	1,272	3	55	,293
Inégalités de résultats	1,804	3	55	,157
Inégalités des chances de résultats	,178	3	55	,911
Inégalités des acquis de base	1,208	3	55	,315

ANOVA						
		Somme des carrés	ddl	Carré moyen	F	Sig.
Efficacité	Intergroupes	,003	3	,001	2,814	,048
	Intragroupes	,018	55	,000		
	Total	,020	58			
Inégalités de résultats	Intergroupes	,002	3	,001	3,227	,029
	Intragroupes	,012	55	,000		
	Total	,014	58			
Inégalités des chances de résultats	Intergroupes	,277	3	,092	5,546	,002
	Intragroupes	,916	55	,017		
	Total	1,193	58			
Inégalités des acquis de base	Intergroupes	,139	3	,046	2,672	,056
	Intragroupes	,952	55	,017		
	Total	1,091	58			

Résultats des tests de comparaisons multiples (Tukey et Bonferroni)								
Variables dépendantes		(I) Types de SABP	(J) Types de SABP	Différence moyenne (I-J)	Erreur standard	Sig.	Intervalle de confiance à 95 %	
							Borne inférieure	Borne supérieure
Efficacité	Tukey	1	2	,01038	,00592	,307	-,0053	,0261
			3	-,00043	,00751	1,000	-,0203	,0195
			4	,01590	,00625	,064	-,0007	,0325
		2	1	-,01038	,00592	,307	-,0261	,0053
			3	-,01081	,00770	,502	-,0312	,0096
			4	,00553	,00648	,829	-,0116	,0227
		3	1	,00043	,00751	1,000	-,0195	,0203
			2	,01081	,00770	,502	-,0096	,0312
			4	,01633	,00796	,182	-,0047	,0374
		4	1	-,01590	,00625	,064	-,0325	,0007
			2	-,00553	,00648	,829	-,0227	,0116
			3	-,01633	,00796	,182	-,0374	,0047
	Bonferroni	1	2	,01038	,00592	,511	-,0058	,0266
			3	-,00043	,00751	1,000	-,0210	,0201
			4	,01590	,00625	,083	-,0012	,0330
		2	1	-,01038	,00592	,511	-,0266	,0058
			3	-,01081	,00770	,996	-,0319	,0103
			4	,00553	,00648	1,000	-,0122	,0233
3	1	,00043	,00751	1,000	-,0201	,0210		

			2	,01081	,00770	,996	-,0103	,0319
			4	,01633	,00796	,269	-,0054	,0381
		4	1	-,01590	,00625	,083	-,0330	,0012
			2	-,00553	,00648	1,000	-,0233	,0122
			3	-,01633	,00796	,269	-,0381	,0054
Inégalité de résultats	Tukey	1	2	,00203	,00484	,975	-,0108	,0149
			3	,01638*	,00614	,048	,0001	,0326
			4	-,00300	,00511	,936	-,0165	,0105
		2	1	-,00203	,00484	,975	-,0149	,0108
			3	,01435	,00629	,115	-,0023	,0310
			4	-,00503	,00530	,778	-,0191	,0090
		3	1	-,01638*	,00614	,048	-,0326	-,0001
			2	-,01435	,00629	,115	-,0310	,0023
			4	-,01938*	,00650	,022	-,0366	-,0022
		4	1	,00300	,00511	,936	-,0105	,0165
			2	,00503	,00530	,778	-,0090	,0191
			3	,01938*	,00650	,022	,0022	,0366
	Bonferroni	1	2	,00203	,00484	1,000	-,0112	,0153
			3	,01638	,00614	,060	-,0004	,0332
			4	-,00300	,00511	1,000	-,0170	,0110
		2	1	-,00203	,00484	1,000	-,0153	,0112
			3	,01435	,00629	,158	-,0029	,0316
			4	-,00503	,00530	1,000	-,0195	,0095
		3	1	-,01638	,00614	,060	-,0332	,0004
			2	-,01435	,00629	,158	-,0316	,0029
			4	-,01938*	,00650	,026	-,0372	-,0016
		4	1	,00300	,00511	1,000	-,0110	,0170
			2	,00503	,00530	1,000	-,0095	,0195
			3	,01938*	,00650	,026	,0016	,0372
Inégalité des chances de résultats	Tukey	1	2	,04086	,04257	,773	-,0719	,1536
			3	,19855*	,05398	,003	,0555	,3416
			4	-,01709	,04497	,981	-,1362	,1020
		2	1	-,04086	,04257	,773	-,1536	,0719
			3	,15769*	,05533	,030	,0111	,3043
			4	-,05795	,04657	,602	-,1813	,0654
		3	1	-,19855*	,05398	,003	-,3416	-,0555
			2	-,15769*	,05533	,030	-,3043	-,0111
			4	-,21564*	,05719	,002	-,3672	-,0641
		4	1	,01709	,04497	,981	-,1020	,1362
			2	,05795	,04657	,602	-,0654	,1813
			3	,21564*	,05719	,002	,0641	,3672
	Bonferroni	1	2	,04086	,04257	1,000	-,0757	,1574
			3	,19855*	,05398	,003	,0508	,3463
			4	-,01709	,04497	1,000	-,1402	,1060
		2	1	-,04086	,04257	1,000	-,1574	,0757
			3	,15769*	,05533	,037	,0063	,3091
			4	-,05795	,04657	1,000	-,1854	,0695
		3	1	-,19855*	,05398	,003	-,3463	-,0508
			2	-,15769*	,05533	,037	-,3091	-,0063
			4	-,21564*	,05719	,002	-,3722	-,0591
		4	1	,01709	,04497	1,000	-,1060	,1402
			2	,05795	,04657	1,000	-,0695	,1854
			3	,21564*	,05719	,002	,0591	,3722

Inégalités des acquis de base	Tukey	1	2	-.04806	,04340	,687	-,1631	,0669
			3	,04886	,05504	,811	-,0970	,1947
			4	-,09839	,04585	,152	-,2199	,0231
		2	1	,04806	,04340	,687	-,0669	,1631
			3	,09692	,05641	,324	-,0525	,2464
			4	-,05033	,04749	,715	-,1761	,0755
		3	1	-,04886	,05504	,811	-,1947	,0970
			2	-,09692	,05641	,324	-,2464	,0525
			4	-,14725	,05831	,067	-,3017	,0072
		4	1	,09839	,04585	,152	-,0231	,2199
			2	,05033	,04749	,715	-,0755	,1761
			3	,14725	,05831	,067	-,0072	,3017
	Bonferroni	1	2	-.04806	,04340	1,000	-,1669	,0707
			3	,04886	,05504	1,000	-,1018	,1995
			4	-,09839	,04585	,218	-,2239	,0271
		2	1	,04806	,04340	1,000	-,0707	,1669
			3	,09692	,05641	,548	-,0575	,2513
			4	-,05033	,04749	1,000	-,1803	,0796
		3	1	-,04886	,05504	1,000	-,1995	,1018
			2	-,09692	,05641	,548	-,2513	,0575
			4	-,14725	,05831	,087	-,3069	,0124
		4	1	,09839	,04585	,218	-,0271	,2239
			2	,05033	,04749	1,000	-,0796	,1803
			3	,14725	,05831	,087	-,0124	,3069

*. La différence moyenne est significative au niveau 0.05.