Mesurer et comparer les compétences des élèves : les leçons de la réforme du PASEC

Hélène Charton

▶ **To cite this version:**

Hélène Charton. Mesurer et comparer les compétences des élèves : les leçons de la réforme du PASEC. Cahiers de la recherche sur l'éducation et les savoirs, 2017, HORS SERIE 5, pp.153-171. halshs-02432421

HAL Id: halshs-02432421
https://shs.hal.science/halshs-02432421

Submitted on 8 Jan 2020

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.

CAHIERS DE LA RECHERCHE SUR L'ÉDUCATION ET LES SAVOIRS

Cahiers de la recherche sur l'éducation et les savoirs

Hors-série n° 5 | 2017
L'école en mouvements

Mesurer et comparer les compétences des élèves : les leçons de la réforme du PASEC

Hélène Charton

OpenEdition Journals

Édition électronique
URL : http://journals.openedition.org/cres/3149
ISSN : 2265-7762

Éditeur
Les éditions de la Maison des sciences de l'Homme

Édition imprimée
Date de publication : 20 décembre 2017
Pagination : 153-171
ISBN : 9782735124138
ISSN : 1635-3544

Ce document vous est offert par Université Bordeaux Montaigne

Université BORDEAUX MONTAIGNE

Référence électronique
Hélène Charton, « Mesurer et comparer les compétences des élèves : les leçons de la réforme du PASEC », *Cahiers de la recherche sur l'éducation et les savoirs* [En ligne], Hors-série n° 5 | 2017, mis en ligne le 29 janvier 2018, consulté le 08 janvier 2020. URL : http://journals.openedition.org/cres/3149

Cahiers de la recherche sur l'éducation et les savoirs est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Mesurer et comparer les compétences des élèves : les leçons de la réforme du PASEC

*Hélène CHARTON**

Cet article porte sur la mesure des acquis des élèves (*learning outcomes*), à partir de l'exemple de la réforme du Programme d'analyse des systèmes éducatifs de la CONFEMEN[1] (PASEC), engagée depuis 2012. Les nouveaux instruments et protocoles d'évaluation adoptés par ce programme, qui suivent désormais les standards internationaux, visent explicitement à mesurer et à comparer les compétences des écoliers d'Afrique francophone. Le changement de rationalité des évaluations standardisées du PASEC, qui avaient d'abord été conçues pour identifier les facteurs d'amélioration de la qualité des apprentissages, invite à réfléchir aux nouveaux rôles et enjeux de l'évaluation dans la gestion et le pilotage des systèmes d'éducatifs.

Caractéristique majeure de la « bureaucratisation du monde » (Hibou, 2012), la question de l'évaluation a pris une importance grandissante au cours des vingt dernières années avec l'augmentation de la demande de chiffres et de données visant à optimiser et améliorer les performances des systèmes éducatifs. La multiplication des évaluations internationales standardisées, de type PISA (*Progress for International Student Assessments* – Programme international pour le suivi des acquis des élèves), témoigne de l'emprise de cette fièvre évaluatrice dans le secteur de l'éducation, indissociable de la généralisation des nouveaux modèles

* Chargée de recherche au CNRS, Les Afriques dans le Monde, Université de Bordeaux

1 Conférence des ministres de l'Éducation des pays africains et malgache d'expression française. Cette institution francophone a été créée en 1960 pour accompagner le développement des systèmes éducatifs des pays nouvellement indépendants. Avec 15 membres à son origine, elle compte aujourd'hui 44 membres et son activité s'étend désormais en dehors de l'Afrique.

de gestion publique et notamment de la gestion axée sur les résultats. La réforme du PASEC s'inscrit dans cette dynamique de production de chiffres et de résultats, qui permettent de définir des seuils et des *benchmarks*, à partir desquels sont élaborés des modèles de bonne gestion.

La plongée dans un espace bureaucratique ordinaire : en l'occurrence ici un séminaire de validation des nouveaux instruments du PASEC, permet de dévoiler les réalités et les enjeux du travail technocratique en action. Les échanges autour des nouveaux instruments et protocoles du programme mettent en évidence la diversité des rationalités qui s'affrontent sur le terrain de l'évaluation et qui sont portées par différentes catégories d'acteurs. À la croisée d'espaces politiques, techniques et internationaux, le PASEC offre une interface pertinente pour rendre compte de la complexité du processus d'élaboration et de négociation de ce type d'instruments d'évaluation. L'analyse des modalités concrètes de sa réforme éclaire, en effet, la portée et le poids des instruments et des raisonnements techniciens dans l'évolution des stratégies d'évaluation et leur impact sur les politiques d'éducation.

Après une brève présentation du PASEC et du contexte dans lequel sa réforme a été initiée, je m'intéresserai plus spécifiquement à l'impact des options techniques retenues dans le processus de réforme sur le sens et la finalité des évaluations et sur la manière dont celles-ci modifient la perception des réalités scolaires en fonction de leur propre logique. Sur la base de cette analyse, il s'agira enfin de tirer les leçons d'une telle réforme et de la place de l'évaluation dans la gestion publique et le processus d'élaboration de normes internationales.

Contexte et enjeux de la réforme du PASEC

Le cadre de l'analyse

Résolument empirique, cette analyse du processus de réforme du PASEC entend contribuer aux débats sur la « fièvre de l'évaluation[1] », en lien avec la généralisation des techniques issues du *New Public Management* (NPM), et notamment la gestion axée sur les résultats, emblématique de la gouvernementalité néolibérale

1 Voir, notamment, le dossier consacré à cette question par la *Revue d'histoire moderne et contemporaine,* n° 55, 2008, intitulé « La fièvre de l'évaluation ».

(Bruno & Didier, 2013 ; Normand, 2005 ; Languille, 2014 ; Dardot & Laval, 2009 ; Laval & Weber, 2002 ; Samuel, 2009). Si la littérature sur le gouvernement par les chiffres (en Afrique et dans les pays en développement) s'est considérablement étoffée au cours des dernières années, l'analyse proposée, qui s'appuie sur ces éclairages, s'intéresse davantage aux techniques de production de résultats qu'aux chiffres eux-mêmes (Desrosières, 2013, 2014 ; Supiot, 2015 ; Fioramonti, 2014 ; Jerven, 2013 ; Ward 2004 ; Cussó, 2010). Il s'agit en effet de centrer la focale sur les techniques et les outils qui nourrissent la vague du NPM, afin de montrer comment leurs modes de fonctionnement transforment les logiques de l'évaluation et impactent la gestion des systèmes éducatifs, et plus généralement les politiques d'éducation.

Sur le plan théorique, cette analyse se réfère à la notion d'instruments, appréhendés comme des institutions articulant les logiques d'acteurs et les référentiels qui structurent les politiques publiques (Lascoumes & Le Galès, 2005). L'emprise grandissante des chiffres et des données statistiques, produits par divers instruments, participe à une forme d'objectivation et de réduction des réalités sociales et politiques, déclinées sous formes de seuils ou de moyennes qui définissent des normes et des bonnes pratiques (Thévenot, 1994 ; Desrosières, 2014 ; Cussó, 2007). L'analyse des lieux et des conditions de production de ce type d'instruments permet de souligner leur contingence en mettant en évidence les compromis et les rapports de pouvoir dont ils sont le produit, comme ici dans le cas de la réforme du PASEC (Hugon, 2013). À l'interface entre différents espaces nationaux, régionaux, internationaux, techniques et pédagogiques mais également politiques, l'analyse de ce programme permet de rendre compte de la complexité et de la diversité des logiques qui sous-tendent la problématique de l'évaluation. Derrière l'hyper rationalité technique généralement véhiculée par ce type d'instruments, l'observation d'une bureaucratie en action donne à voir les conflits et les luttes engagés sur le terrain des méthodes et des techniques qui mobilisent un répertoire varié de références et de représentations.

Sur le plan méthodologique, l'article s'appuie sur une enquête de terrain menée auprès du PASEC entre novembre 2012 et avril 2014 à Dakar. L'observation d'un séminaire technique organisé à Saly sur la Petite Côte au Sénégal du 18 au 28 février 2013 pour présenter, valider et discuter les nouveaux instruments d'évaluation du programme constitue l'un des moments clés de l'enquête. Animé par l'équipe technique du PASEC, composée de huit conseillers techniques, ce séminaire réunissait quatre représentants de chacune des 10 équipes nationales participant à la première vague d'évaluations. Les délégations nationales

étaient composées pour moitié de statisticiens et d'inspecteurs pédagogiques. Outre l'observation du séminaire, structuré autour de présentations et de discussions sur les nouveaux instruments et protocoles d'évaluation, plusieurs entretiens individuels ont été menés auprès de membres des équipes nationales et de l'équipe technique du PASEC basée à Dakar. Un travail important d'analyse et de synthèse a été réalisé à partir des documents discutés et produits pendant l'atelier[2]. Ces données originales, croisées avec la documentation interne du PASEC (rapports d'évaluation, notes) permettent d'appréhender dans leur diversité et leur complexité les enjeux soulevés par la réforme en lien avec le positionnement des différents acteurs impliqués dans ce processus.

Le PASEC et l'évaluation des acquis des élèves

Le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) a été créé en 1991 par la coopération française, dans le sillage de la conférence de Jomtien en 1990 sur l'Éducation pour tous. Ses objectifs étaient les suivants : identifier des modèles d'écoles efficaces et peu coûteux ; développer dans chacun des États participants, une capacité interne et permanente d'évaluation de son système éducatif ; diffuser largement les méthodes, instruments et résultats en matière d'évaluation. Les évaluations du PASEC avaient adopté la méthode de la « valeur ajoutée », mise au point par des chercheurs de l'IREDU (Institut de recherche sur l'éducation de l'Université de Bourgogne) visant à isoler les facteurs pédagogiques et sociaux-économiques susceptibles d'améliorer la qualité des apprentissages. Des tests étaient réalisés sur un échantillon représentatif d'élèves de 2e et de 5e année de primaire, en début et en fin d'année scolaire. Ces résultats étaient ensuite corrélés par une méthode d'analyse multivariée, avec des données relatives à l'environnement scolaire et familial recueillies auprès des élèves, des maîtres et des directeurs d'écoles.

En tant que programme de la CONFEMEN, le PASEC est piloté par ses instances exécutives, à savoir le Secrétaire général et les sessions ministérielles annuelles. Produit de l'aide bilatérale, l'Agence française de développement

2 Je remercie la CONFEMEN de m'avoir permis d'assister à cette réunion de façon indépendante et d'avoir eu accès à ces informations. Je remercie également l'équipe technique du PASEC et les membres des équipes nationales pour leur confiance et leur disponibilité dans l'exercice de l'enquête.

(AFD) est le principal bailleur extérieur du programme, dont elle assure le financement sur la base d'un contrat triennal reconduit en fonction de ses résultats. La coordination technique du programme est confiée à un expert recruté et payé par la coopération française au titre de l'assistance technique. Il organise les activités des huit conseillers techniques, répartis entre les divisions « instruments et procédures d'enquête » et « gestion des données et analyse statistique ». Leurs profils et parcours sont assez hétérogènes : plutôt jeunes (moins de quarante ans), certains ont poursuivi au sein de l'organisation une activité commencée comme volontaire international, d'autres sont passés par des programmes internationaux ou sont issus des services techniques (planification statistique) des ministères africains de l'Éducation nationale. Ils sont chargés de l'organisation des campagnes nationales d'évaluation et du traitement des données en lien avec les équipes nationales PASEC qui ont été progressivement mises en place à la demande des gouvernements. Ces équipes sont constituées de cadres des ministères de l'Éducation des pays concernés et relèvent à la fois des services pédagogiques (inspecteurs) et techniques (statisticiens).

Alors que le PASEC avait conduit en moyenne 6 à 8 évaluations annuelles entre 1991 et 2005, ses activités se sont ensuite amplifiées et diversifiées avec la réalisation de 18 enquêtes pays et de plusieurs évaluations thématiques entre 2006 et 2010. Cette recrudescence d'activités s'est traduite par une augmentation significative de son personnel, initialement composé de deux conseillers techniques. À partir de 2005, l'équipe comptait en plus un volontaire international et deux postes temporaires, qui ont ensuite été pérennisés, pour des experts des équipes nationales, ainsi que la création d'un poste de statisticien et de conseiller en communication, avant de connaître la configuration actuelle. L'étoffement de l'équipe est directement lié aux évolutions des contextes à la fois politiques et institutionnels dans lesquels évolue le PASEC. La mobilisation générale en faveur de l'éducation primaire universelle, deuxième objectif du millénaire pour le développement (OMD), a stimulé l'émergence d'une ingénierie éducative spécifiquement dédiée à la réalisation de ces objectifs. C'est dans cette logique que le Pôle de Dakar a été créé en 2001 par la coopération française dans le but de renforcer l'analyse sectorielle en éducation, à travers la mise en place d'outils diagnostics comme les Rapport sur l'état du système éducatif national (RESEN). Ces dispositifs sont étroitement liés, mobilisant des techniques, références et équipes identiques. Ce sont, en effet, les mêmes économistes de l'IREDU, notamment Alain Minguat et Paul Coustère, à l'origine du PASEC, qui ont développé les RESEN, le Pôle de Dakar et tissé des liens étroits

avec la Banque mondiale dans le cadre de la mise en place de l'Initiative Fast Track (Samuel, 2013).

La réforme du PASEC dans son contexte

Ces différents programmes répondent à la demande constante de résultats et de chiffres générée par les instruments de gestion de l'aide internationale appliqués aux systèmes éducatifs, que les politiques de renforcement de l'efficacité de l'aide ont encore accentuée, depuis la Déclaration de Paris en 2004. La multiplication des programmes standardisés d'évaluation des acquis des élèves, encouragés par les institutions internationales dans les pays en développement, s'inscrit dans cette même logique.

Outre les programmes régionaux dédiés à l'Afrique, comme le PASEC ou son équivalent pour les pays africains anglophones, le SACMEQ (*Southern and Eastern Africa Consortium for Monitoring Educational Quality*) créé en 1995 en partenariat avec l'Unesco[3], les enquêtes internationales TIMSS (*Trends in International Mathematics ans Science Study*) et PIRLS (*Progress of International Reading Litteracy Study*) ont élargi leur champ d'activité en direction des pays en développement. Le programme PISA de l'OCDE (Organisation de coopération et de développement économiques), qui cible les compétences des jeunes de 15 ans, vient également de lancer *PISA for Development*, (OCDE, 2013). Enfin des programmes d'évaluation plus spécifiques ont vu le jour comme les tests EGMA (*Early Grades Assessments in Mathematics*) et EGRA (*Early Grades Assessments in Litteracy*) mis au point par la Banque mondiale et l'agence américaine d'aide bilatérale – USAID – pour mesurer les apprentissages précoces des élèves en début de primaire en mathématiques et en littéracie. Entre 2006 et 2011, 41 évaluations de ce type ont été conduites dans les pays en développement (Bourdon & Altinok, 2012). Alors que le PASEC occupait une position de quasi-monopole sur l'évaluation des acquis des élèves dans les pays francophones d'Afrique de l'Ouest, l'émergence de nouveaux acteurs dans cet espace mais surtout les nouveaux enjeux de l'évaluation des acquis des élèves dans le

3 Les deux programmes régionaux n'ont originellement aucun lien institutionnel, c'est seulement dans le cadre de la réforme décrite qu'un rapprochement a été amorcé en vue d'homogénéiser les modes d'évaluations, jusque-là fort différents.

contexte de la généralisation de la gestion par les résultats des systèmes éducatifs ont amené le programme à se repositionner.

La question de la place du PASEC dans le paysage général des évaluations internationales a été abordée pour la première fois en 2004, lors de la 51e session ministérielle de la CONFEMEN, consacrée à l'évaluation. Le rapport commandé alors, publié en 2006 sous le titre, « L'évaluation des systèmes éducatifs : pour un meilleur pilotage par les résultats » avait souligné la faiblesse des dispositifs d'évaluation existants et la nécessité de renforcer leur articulation avec les programmes internationaux d'évaluation. La méthodologie utilisée par le PASEC avait fait l'objet de nombreuses critiques soulevées lors des évaluations externes du programme, pointant son manque de robustesse (en matière d'échantillonnage notamment), en comparaison des modèles psychométriques d'évaluation, comme ceux adoptés par PISA considéré comme une référence mondiale en matière de tests standardisés. Les anciens tests PASEC mesuraient les progrès réalisés par les élèves entre deux moments et la corrélation de ces résultats avec des indicateurs socio-éducatifs permettait d'identifier les facteurs d'amélioration de la qualité des apprentissages. Environ 25 variables qualitatives ont ainsi été isolées mais l'impact global de ces évaluations s'est révélé assez faible. La lourdeur du dispositif d'enquête en termes de coût (plus de € 100 000) et de durée (trois ans en moyenne) limitait le nombre d'enquêtes réalisées dans chaque pays, et n'a finalement pas permis de mettre en évidence l'existence d'un profil type « d'écoles efficaces et peu coûteuses » (CRES, 2011). Sur la base de ce constat, la CONFEMEN avait préconisé, en 2008, la réorientation de son programme d'évaluation : « pour assurer la crédibilité nécessaire à l'évolution du programme, celui-ci doit développer les compétences en psychométrie afin d'élaborer de nouveaux tests, conformes aux standards internationaux et d'assurer une appropriation et une utilisation efficace de ces nouveaux instruments » (CONFEMEN, 2008). C'est dans ce contexte et pour répondre à ces exigences, que l'équipe technique du PASEC a entrepris la révision de ses instruments et protocoles d'évaluation.

De nouveaux instruments au service de la mesure et de la comparaison des compétences des élèves

La refonte des instruments d'évaluation, qui est le principal objet de cette réforme, a des implications qui dépassent largement ses seuls enjeux techniques.

Le choix des méthodes et des outils modifie, en effet, le sens et la finalité de ces évaluations mais également la perception des réalités scolaires.

Comparer des compétences

Afin de pouvoir mesurer et comparer les compétences des élèves du primaire en mathématiques et en littéracie, la méthodologie distinctive de la valeur ajoutée a été abandonnée au profit de tests psychométriques. Les tests psychométriques de type MRI[4] (Modèles de réponses aux items), déjà utilisés par les programmes PISA, PIRLS, TIMMS et SACMEQ, poursuivent des objectifs radicalement différents puisqu'ils visent à mesurer le degré de compétence des élèves par rapport à une tâche spécifique (Vrignaud, 2006). L'identification de seuils de compétence par rapport à une échelle de mesure définie par le test favorise la comparaison et le *benchmark*, qui sont l'un des objectifs de ce type d'évaluation.

Sur un plan technique, le passage à la psychométrie a nécessité de réviser les procédures de construction des items (questions posées), les modalités d'échantillonnage et de passation des tests. La mesure des compétences par la méthode MRI repose sur un certain nombre de règles et principes. Ainsi, les réponses aux items doivent être indépendantes les unes des autres (les réponses ne doivent pas dépendre de réponses antérieures) et les questions doivent être formulées de manière à isoler une seule variable, rendant compte à la fois de la compétence de l'élève et de la difficulté de l'item (Vrignaud, 2006). Enfin, les tests doivent pouvoir mesurer des compétences avérées, définies au préalable. Concrètement, cela s'est traduit, au niveau du PASEC, par la mise en concordance des tables de spécification qui font le lien entre les curricula, les items et les compétences testées. L'équipe du PASEC a ainsi réalisé, en lien avec plusieurs équipes de consultants, un travail important d'analyse des différents curricula des pays testés. Son objectif, à moyen terme, est de produire « des échelles de compétences comparables en compréhension de l'écrit et en mathématiques ainsi qu'une batterie d'indicateurs contextuels communs à l'échelle du continent africain en fin de scolarité primaire (6e année) » (PASEC, 2012).

4 Le modèle MRI est un modèle mathématique permettant de représenter la relation entre la compétence de l'élève (définie à partir du niveau de discrimination de l'item) et la difficulté de l'item (évaluée en fonction de la proportion d'élèves le réussissant).

Cette ambition continentale a mené au rapprochement du PASEC avec le programme régional anglophone. Les nouvelles évaluations sont désormais réalisées par vagues d'une dizaine de pays, à l'instar de ce qui se pratique au niveau du SACMEQ et suivant un calendrier commun. Mais pour que les données recueillies par les deux programmes soient comparables, il est nécessaire d'étalonner les résultats des élèves testés en définissant un nombre significatif de questions communes (items d'ancrage). Enfin, les élèves sont désormais testés en 6e année de primaire (qui correspond à la fin du cycle dans la plupart des pays du continent) et non plus en 5e année.

La dimension résolument comparative des nouvelles évaluations PASEC et l'adoption des normes psychométriques affectent également les modalités de passation des tests. Pour assurer l'indépendance des réponses et éviter que les élèves puissent revenir en arrière en cours d'exercice, de nouveaux cahiers d'évaluations, cachetables au fur et à mesure des épreuves, ont été fabriqués. Les tests de deuxième année de primaire ont également été modifiés. Leur finalité est moins la comparaison que l'évaluation de la qualité des apprentissages précoces sur le modèle des tests EGRA et EGMA, dont ils reprennent certaines modalités. Ainsi les tests ne sont plus auto-administrés mais réalisés par des administrateurs qui interrogent les élèves. De nouvelles techniques de mesures, inspirées des EGRA et EGMA, ont aussi été introduites. Afin de mesurer les compétences en lecture, on compte par exemple combien de mots un élève est capable de lire en un temps donné. Ces épreuves chronométrées ont nécessité de faire fabriquer des minuteurs spécifiques, faciles à manipuler et à dissimuler aux élèves. Pour contrebalancer le surcoût lié à ce type de passation, la taille de l'échantillon a été réduite : 12 élèves (au lieu de 20) testés dans 90 classes en 2e année et toujours 20 élèves dans 180 classes de 180 écoles en 6e année.

Au total, l'adoption des nouveaux protocoles psychométriques d'évaluation par le PASEC affecte aussi bien la conception des tests que leur dimension matérielle à laquelle les responsables du programme accordent la plus grande attention. Cette dernière participe, en effet, pleinement du succès de la réforme, et par là de la reconnaissance et de la légitimité internationale du nouveau programme. Ce détour par la fabrique matérielle des évaluations internationales donne à voir l'emprise des logiques techniciennes, inhérentes au choix des instruments.

« La production bureaucratique du réel » (Samuel, 2013)

Les nouveaux protocoles d'évaluation du PASEC ont été présentés et discutés avec les représentants des équipes nationales participant à la première vague de tests groupés lors d'un séminaire régional organisé à Saly, au Sénégal, en février 2013. Ces instruments, orientés vers la production de comparaisons ne modifient pas seulement le sens et la finalité des évaluations PASEC, mais également la manière dont sont appréhendées les réalités scolaires. L'observation du séminaire de Saly et les discussions relatives à la construction et à la formulation des items ou aux modalités de passation et d'échantillonnage, mettent en évidence les conflits et les tensions entre les rationalités portées par les acteurs en présence. Les clivages entre les concepteurs des dispositifs d'évaluations construits sur la base de références techniques exogènes et les équipes chargées de les mettre en œuvre sur le terrain, révèlent la distance qui sépare les réalités scolaires perçues et vécues par leurs praticiens de celles construites par la rationalité statistique. Les évaluations standardisées doivent en effet pouvoir comparer des espaces scolaires statistiquement comparables, tant au niveau de leur représentativité que des compétences testées. Par conséquent, les instruments d'évaluation ne rendent jamais compte des réalités scolaires telles qu'elles existent mais relèvent d'une construction produite par une série de réductions, à la fois statistiques (échantillonnage) et pédagogiques (compétences mesurées). Les évaluations orientées vers la production de comparaisons tendent à renforcer les distorsions avec les réalités locales en écrasant les différences. Ce décalage entre les données recueillies par des instruments à mailles larges, conçus pour produire de la comparaison, et la réalité singulière des systèmes nationaux d'éducation, est clairement ressorti des discussions menées à Saly, comme l'illustrent les exemples suivants.

Le premier exemple porte sur l'échantillonnage et la représentativité des écoles testées par rapport aux réalités scolaires nationales. Pour des raisons pratiques et matérielles (langues d'enseignement), les écoles dont la langue d'enseignement est l'arabe ont été exclues des échantillons du PASEC, bien qu'elles scolarisent un nombre important d'élèves dans les pays d'Afrique de l'Ouest. Alors que les écoles communautaires[5] représentent un phénomène social

5 Les écoles communautaires désignent en Afrique de l'Ouest des espaces d'apprentissages (classes, écoles) ouverts et pris en charge par les communautés (construction d'un abri ou des bâtiments, recrutement et rétribution de l'enseignant) pour répondre aux carences de l'offre éducative.

et politique important dans certains pays comme le Tchad, ces écoles ne font pas non plus partie de l'échantillon, en raison du caractère lacunaire des données dont disposent les administrations à leur sujet. Comme l'ont expliqué les responsables du PASEC aux membres des équipes nationales qui avaient soulevé cette question, les arbitrages en matière d'échantillonnage, et par conséquent leur degré de finesse et de précision, dépendent des moyens investis pour les réaliser, et donc du budget que chaque pays est prêt à consacrer à ce type d'enquête. Dans cet exemple, la dimension très politique de la représentativité des espaces scolaires est occultée par des justifications techniques et économiques qui renvoient les États à leurs propres responsabilités.

Le second exemple concerne les modalités pratiques de passation des tests et la prise en compte des réalités d'apprentissage des élèves. À l'exception du Burundi, où l'enseignement primaire se fait en kirundi, le français est partout en Afrique francophone la langue officielle d'enseignement et c'est dans cette langue que sont réalisés les tests PASEC. Dans les petites classes, la maîtrise du français est cependant loin d'être acquise et jamais homogène. Pour des raisons pratiques, liées notamment au recrutement et à la formation des administrateurs des tests de 2e année et afin de garantir des conditions homogènes de passation, les administrateurs ne peuvent pas aider les élèves en traduisant par exemple la consigne en langue locale, comme cela se fait *a priori* dans la plupart des situations réelles d'enseignement. Des informations relatives à l'usage du français hors de l'école sont cependant recueillies grâce à des questionnaires adressés aux élèves et aux enseignants, pour permettre de pondérer les résultats en fonction du degré d'exposition à la langue d'enseignement. Il n'empêche que les modalités d'administration des tests créent un hiatus avec les situations réelles d'apprentissage et posent directement la question de la compétence effectivement évaluée dans les cas nombreux où l'enfant ne maîtrise pas la langue dans laquelle sont passés les tests.

La prise en compte des pratiques et des réalités effectives d'apprentissage dans ce type d'évaluation est, en effet, loin d'être évidente. Les questions adressées aux enseignants pour évaluer, par exemple, le temps alloué aux différents apprentissages ou encore le degré d'application de certaines méthodes pédagogiques comme l'approche par les compétences (APC), afin d'apprécier la réalité des pratiques enseignantes par rapport aux programmes officiels, ont fait l'objet de vives critiques de la part de membres des équipes nationales. Bien que plusieurs études aient souligné l'application limitée de pratiques pédagogiques comme l'APC dans les salles de classe, les représentants des administrations scolaires ont

systématiquement affirmé que les pratiques pédagogiques officiellement actées, comme l'APC, étaient bien ancrées dans les classes. De leur point de vue, les instruments d'évaluation doivent valider les prescriptions officielles, même si elles sont peu ou pas appliquées sur le terrain et même si elles biaisent la mesure des compétences effectives des élèves. La reconnaissance du bricolage des prescriptions auquel doivent se livrer nombre d'instituteurs pour les ajuster aux réalités de leur salle de classe ne remet pas seulement en cause l'ordre scolaire officiel mais affecte directement leur légitimité et par conséquent leur autorité. Dans les discussions techniques qui leur échappaient, les pédagogues (inspecteurs et cadres pédagogiques) ont toujours cherché à affirmer la légitimité et l'autorité de leur savoir et de leur connaissance intime de l'école et des processus d'apprentissage face à une vision technicienne de l'école, notamment défendue par les statisticiens. Cette rationalité bureaucratique de l'administration scolaire, qui s'oppose cette fois à la raison statistique, introduit une nouvelle distorsion dans l'appréhension des réalités scolaires.

Au total, ces exemples révèlent les différentes perceptions des réalités scolaires qui relèvent de rationalités variées et qui ne sont jamais données mais construites. Si les instruments, qui opèrent au sein d'un espace techniquement contraint, modèlent et projettent une certaine vision de l'école, ils doivent composer avec les perceptions des agents chargés de les mettre en œuvre. Au cours de cet exercice de validation, les conseillers se sont efforcés de maintenir la discussion à un niveau strictement technique. Ainsi les questions pourtant soulevées, relatives au sens et à la pertinence de cette réforme, dans des contextes nationaux où les capacités internes d'évaluation demandent partout à être développées, sont demeurées sans réponse. Ce hiatus illustre finalement la tension récurrente entre l'universel et le particulier qui est au cœur de la question de la mesure et de l'évaluation dont l'une des fonctions est, précisément, de « dépasser la subjectivité et le particularisme de l'observateur et de l'observé à travers la mobilisation d'outils qui agissent précisément comme des médiateurs entre les champs théoriques et empiriques » (Hugon, 2013 : 105). Tout l'enjeu du séminaire de Saly était précisément de dépasser ces clivages et susciter l'adhésion de l'ensemble des acteurs à ces instruments.

Gouverner par les chiffres : Les nouveaux enjeux de l'évaluation

Les modalités concrètes du processus de réforme du PASEC ont mis en évidence la part des différentes logiques techniciennes et institutionnelles qui président à la production d'instruments d'évaluation dont le périmètre d'action ne correspond pas à l'espace concret de fonctionnement des écoles. Le changement de rationalité des évaluations PASEC dont l'objectif premier n'est plus d'identifier les facteurs d'amélioration de la qualité de l'éducation mais de comparer à l'échelle régionale les compétences des élèves de dernière année de primaire, met en évidence le nouveau rôle de l'évaluation des acquis des élèves dans la gestion des systèmes éducatifs africains.

Les comparaisons, qui sont au cœur de la réforme du PASEC, permettent de classer et de déterminer des seuils qui servent à repérer des bonnes pratiques sur la base desquelles sont élaborées les politiques publiques. Dans un contexte où la recherche d'efficacité et la bonne gestion sont un principe moteur de la régulation et de la gestion des systèmes éducatifs, ces instruments d'évaluation, pourvoyeurs de chiffres, ont un réel pouvoir prescriptif. Les résultats des évaluations standardisées sont destinés en premier lieu aux bailleurs et aux décideurs, soucieux de pouvoir contrôler l'efficacité de leur aide et de leurs actions sur la base de résultats qu'ils jugent crédibles. Ce sont précisément ces chiffres, au cœur de l'ingénierie réformatrice déployée par les bailleurs envers les États africains, qui servent à donner la température de l'école lors des revues sectorielles annuelles, ou encore à élaborer des requêtes de financement externe et à construire les indicateurs conditionnant la mise en œuvre de réformes et la conduite de programmes. Or pour être reconnus comme valides et légitimes, les modalités de production de ces chiffres doivent être endossées par l'ensemble des acteurs. Cette dernière dimension permet précisément de mettre en évidence les enjeux d'une réforme comme celle du PASEC dans la fabrique et la négociation de normes d'évaluation autour desquelles se structure le gouvernement de l'école.

Comparer pour « bien gouverner » : les leçons d'une réforme

Les nouveaux tests du PASEC visent moins à améliorer les conditions d'apprentissage des élèves qu'à mesurer et comparer les performances des systèmes éducatifs dans un souci de bonne gestion. En établissant un lien entre ce que les élèves apprennent à l'école et l'usage potentiel de ces apprentissages dans le monde

professionnel, les évaluations psychométriques fournissent une certaine mesure de l'efficacité des systèmes éducatifs (Vrignaud, 2006). En outre, la production de données comparables offre la possibilité de classer les différents systèmes éducatifs en fonction des performances de leurs élèves et ainsi identifier de bonnes pratiques et des modèles de réussite suivant la technique éprouvée du *benchmarking*, instrument classique de la production de politiques publiques internationales dans un contexte libéral. On peut, sous cet angle, lire la réforme du PASEC comme une émanation de l'ingénierie réformatrice issue du *New Public Management*, « puzzle doctrinal » dont la gestion par les résultats et la mesure des performances constituent certains des traits saillants (Bèzes, 2007).

Du point de vue de ses promoteurs, l'un des enjeux de la réforme du programme est précisément de contribuer au renforcement et à la diffusion auprès des acteurs de l'éducation d'une culture de l'évaluation et de la gestion par les résultats dans le but explicite d'améliorer la gouvernance de l'école. C'est en effet dans leur stratégie de renforcement des capacités des États africains que s'inscrit l'appui des bailleurs internationaux à la réforme du PASEC. En 2005, la Banque mondiale avait initié un programme d'appui aux structures d'évaluation nationales et régionales comme le PASEC et le SACMEQ. Le partenariat technique et financier conclu avec le PASEC autour d'un programme de renforcement des capacités en évaluation des pays membres avait débouché en 2012 sur la création d'un « Réseau régional de l'Afrique subsaharienne francophone pour la Gouvernance axée sur les résultats dans le secteur de l'Éducation ». L'un des principaux objectifs de ce programme de trois ans, financé à hauteur de US$ 1M, était de renforcer les capacités en gestion axée sur les résultats des analystes et des responsables politiques liés à la gestion des systèmes éducatifs des administrations nationales. Dans ce cadre, plusieurs ateliers de formation internationaux qui leur étaient spécifiquement dédiés ont été organisés pour renforcer leur sensibilisation et leur l'implication sur les questions d'évaluation. Ils étaient animés par les conseillers techniques du PASEC et des consultants extérieurs. Des ressources spécifiques ont également été allouées pour former et équiper les membres des équipes nationales et les conseillers techniques du PASEC aux approches psychométriques. La technique et les instruments, ainsi mis au service des stratégies de renforcement des capacités des administrations africaines, véhiculent de nouvelles pratiques évaluatives indissociables des modes spécifiques de gestion par les résultats. L'adhésion de l'ensemble des acteurs de l'éducation (praticiens décideurs, gestionnaires) est au cœur de ce processus de réforme qui a une dimension éminemment politique.

Malgré les nombreuses formations et ateliers organisés par le PASEC depuis sa création pour renforcer les capacités nationales d'évaluation, leur impact s'est avéré assez limité. L'étude commandée en 2009 par la CONFEMEN pour évaluer la prise en compte et l'impact des résultats des évaluations menées jusque-là par le PASEC avait souligné la faible mobilisation des enquêtes PASEC par les administrations scolaires nationales, qui fustigeaient le plus souvent des résultats et analyses mettant en évidence leurs défaillances (CRES, 2011). Le nouveau dispositif d'évaluation du PASEC vise explicitement à renforcer, à tous les niveaux, l'implication et l'adhésion des agents nationaux. Ainsi, les équipes nationales sont désormais chargées de la rédaction des rapports pays à partir des données recueillies dans le cadre des évaluations groupées. Lors du séminaire de Saly, l'implication des membres de ces équipes dans le processus de validation des instruments visait explicitement à recueillir leur adhésion et leur consentement, afin de faire taire les critiques relatives à la pertinence des outils d'évaluation du PASEC. Les conseillers techniques du PASEC ont, en effet, beaucoup insisté sur le caractère inédit et novateur de cette démarche participative de validation. À un niveau plus politique, l'adhésion passe par l'engagement financier des États, dont la contribution directe aux enquêtes menées par le PASEC est désormais fixée à 35 % de leur coût contre 5 % auparavant.

La production d'une norme internationale et les dessous du consensus

Malgré ces mesures volontaristes, l'adoption des normes psychométriques d'évaluation ne va pas de soi. Les résistances et les stratégies affichées par certains acteurs, en l'occurrence ici certains conseillers techniques du PASEC que nous avons interrogés, éclairent la manière dont s'imposent de nouvelles normes internationales et se fabriquent les consensus.

Le « petit club fermé » des détenteurs du patrimoine du PASEC incarné par ses conseillers techniques, s'est montré peu enclin à abandonner le capital et le savoir-faire technique, sur lequel ils avaient construits leur carrière, au profit des modèles psychométriques dont ils ne sont pas des spécialistes. Plusieurs ont ainsi exprimé des réserves quant à la pertinence et la validité des modèles psychométriques d'évaluation qui relèvent, pour reprendre le propos de l'un d'entre eux, « d'un véritable acte de foi ». Elaborées par un groupe très restreint d'experts (à l'instar de l'ancienne méthodologie PASEC), ces méthodes se sont imposées sur le marché international de l'évaluation, moins en raison de leur robustesse ou de

leur infaillibilité que parce qu'elles répondaient à un besoin spécifique : la mesure et la comparaison des compétences des élèves. L'attitude ambivalente de certains conseillers techniques qui adoptent ces nouveaux instruments malgré leur défiance rend compte de la part des stratégies individuelles et corporatistes dans ces processus d'adhésion qui contribuent à la fabrique silencieuse du consensus (Charton, 2015). Cheville ouvrière de la réforme, les conseillers techniques du PASEC ont tout intérêt à préserver et augmenter leur capital technique en s'appropriant ces nouveaux outils afin de garantir leur survie institutionnelle et défendre leur position hégémonique au sein du programme régional.

La place du PASEC au sein de la CONFEMEN et ses relations avec le Secrétariat général de la CONFEMEN font, depuis plusieurs années, l'objet de tensions qui avaient notamment conduit à la démission du coordonnateur technique en 2010 et à la vacance de ce poste jusqu'en 2013. Le Secrétariat général avait en effet tendance à s'immiscer et à chercher à contrôler les activités du PASEC, considéré comme le programme phare de la CONFEMEN, au nom de logiques politiques compromettant la cohérence et la qualité technique des évaluations. Jalouse de son autonomie et forte de la supériorité technique de son expertise des systèmes éducatifs, l'équipe du PASEC dénonçait ces formes d'instrumentalisation du programme. Ces problèmes de gouvernance ont fait l'objet de l'attention de ses bailleurs et notamment de l'AFD, qui avait commandé, en 2011, une évaluation externe sur « le pilotage, le fonctionnement et les modes de gestion du PASEC dans la perspective de son évolution vers la mise en œuvre d'un dispositif d'évaluations internationales », en vue du renouvellement de son concours pour la période 2012-2016 (AFD, 2012). L'autonomisation du PASEC par rapport à la CONFEMEN avait été envisagée comme un scénario possible de sortie de crise. Mais dans le contexte de la réforme et de l'internationalisation du programme, la survie du PASEC, et de son équipe, passe par le renforcement de son expertise technique. Il s'agit concrètement de s'adapter aux mutations contextuelles et institutionnelles en endossant les nouveaux standards internationaux de l'évaluation qui sortent de son champ d'expertise traditionnelle. La détermination et l'énergie déployées par l'équipe technique du PASEC à faire aboutir une réforme dont elle n'était pas *a priori* partie prenante, est essentiellement motivée par ces stratégies individuelles et collectives de survie et de légitimation du programme francophone dans le champ concurrentiel de l'évaluation internationale.

Conclusion

Au total, la réforme du PASEC définit un nouvel espace de production des normes de l'évaluation scolaire. Endossée par les ministres de l'Éducation réunis au sein de la CONFEMEN et prise en charge par une équipe technique soucieuse d'assurer sa survie dans le champ concurrentiel de l'évaluation internationale, la finalité de cette réforme est moins l'amélioration des apprentissages des élèves que la "bonne gestion" des systèmes éducatifs. Son analyse éclaire les logiques complexes de la fabrique de l'action publique dans un environnement institutionnel international et le rôle de médiateur joué par les instruments. Fruit de compromis et d'arbitrages entre les rationalités statistiques, gestionnaires ou institutionnelles, l'élaboration d'un consensus autour d'instruments sophistiqués, orientés vers la production de comparaisons et l'évaluation des performances des systèmes éducatifs à des fins de régulation, dessine un espace normatif unificateur. Cette plongée dans le fonctionnement d'une bureaucratie en action met en relief le poids des logiques techniciennes, mises au service de la bonne gestion des systèmes éducatifs africains. La réforme du PASEC, permet, en effet, de tirer des leçons sur la manière dont s'imposent certaines normes internationales. Ni meilleures ni pires que d'autres, elles répondent aux besoins et aux stratégies d'acteurs hétérogènes dont les intérêts pourtant divergents les poussent à endosser des normes qui leur permettent de renforcer leur emprise hégémonique dans un domaine spécifique. Dans le domaine de l'évaluation, comme dans beaucoup d'autres, il n'y a jamais *a priori* d'unanimité sur les objectifs et les techniques qui relèvent toujours d'une construction et d'arbitrages entre les acteurs dominants qui défendent légitimement leurs prérogatives et leur position dans un espace donné. À travers la réforme des instruments et des protocoles d'évaluation du PASEC, le positionnement différencié des acteurs qui interviennent dans ce champ en mobilisant un large éventail de répertoires techniques, gestionnaires et institutionnels, révèle la complexité du processus d'émergence et de diffusion de normes internationales.

Bibliographie

Afd, 2012, « Évaluation externe du programme PASEC de la CONFEMEN », Paris, AFD.

Bezès (P.), 2007, « Construire des bureaucraties post-Wébériennes à l'ère du New Public Management », *Critique internationale,* n° 35, pp. 9-29.

Bourdon (B.) et Altinok (N.), 2012, « Renforcer la mesure sur la qualité de l'éducation », Paris, AFD.

Bruno (I.) et Didier (E.), 2013, *Benchmarking. L'État sous pression statistique*. Paris, La Découverte.

Charton (H.), 2015, « Gouvernement de l'école et communautés épistémiques au Sénégal. Trajectoires d'acteurs et circulation des normes et des modèles d'éducation en contexte international », *Revue Tiers Monde*, n° 223, pp. 49-65.

Confemen, 2008, « L'évaluation des systèmes éducatifs : pour un meilleur pilotage par les résultats », Dakar, CONFEMEN.

Consortium pour la recherche économique et sociale (cres), 2011, « Étude sur la prise en compte et l'impact des résultats issus des évaluations PASEC », Dakar, CONFEMEN.

Cussò (R.), 2007, « SAQMEC, PASEC, PISA… les enquêtes sur les acquis des élèves et les politiques de mondialisation », *in* F. Cros et A. Bon, dir., *Les évaluations au niveau international : impacts, contradictions et incertitudes*, Sèvres, CIEP, pp. 79-89.

Cussò (R.), 2010, « Statistique et politique : la force et les limites des données. Des indicateurs-objectif aux indicateurs-condition en éducation », *Éducation Comparée*, n° 3, pp. 45-68.

Dardot (P.) et Laval (C.), 2009, *La nouvelle raison du monde. Essai sur la société néolibérale*. Paris, La Découverte.

Desrosières (A.), 2013, « La mesure du développement : un domaine propice à l'innovation méthodologique », *Revue Tiers Monde*, n° 213, pp. 23-32.

Desrosières (A.), 2014, *Prouver et gouverner*, Paris, La Découverte.

Fioramonti (L.), 2014, *How Numbers Rule the World: the Use and Abuse of Statistics in Global Politics,* Londres, Zed Books.

Hibou (B.), 2012, *La bureaucratisation du monde à l'ère néolibérale*, Paris, La Découverte.

Hugon (P.), 2013, « La mesure du développement est-elle l'expression de rapports de pouvoir ? Application au champ de l'éducation », *Revue Tiers Monde*, n° 213, pp. 105-121.

Jerven (M.), 2013, *Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It*, Ithaca, Cornell University Press.

Lascoumes (P.) et Le Galès (P.) (dir.), 2005, *Gouverner par les instruments*. Paris, Presses de Science Po.

Languille (S.), 2014, « Education through performativity. Learning crisis and technology of quantification in Tanzania », *International Journal of Educational Development,* n° 39, pp. 49-58.

Laval (C.) et Weber (L.), 2002, *Le Nouvel Ordre éducatif mondial. OMC, Banque Mondiale, OCDE, Commission européenne*, Paris, Éditions Nouveaux regards/Syllepse.

Normand (R.), 2005, « La mesure de l'école : politique des standards et management par la qualité », *Cahiers de la recherche sur l'éducation et les savoirs*, HS n° 1, pp. 67-82.

Ocde, 2013, « Improving learning outcomes worldwide: How PISA can help », Paris, OCDE.

Pasec, 2012, « Présentation du partenariat PASEC/SACMEQ », note interne, Dakar, CONFEMEN.

Samuel (B.), 2009, « Le cadre stratégique de lutte contre la pauvreté et les trajectoires de la planification au Burkina Faso », *Sociétés politiques comparées*, n° 16, [en ligne], <http://www.fasopo.org>.

Samuel (B.), 2013, « L'"Éducation pour tous" au Burkina Faso. Une production bureaucratique du réel » *in* B. Hibou, dir., *La bureaucratisation néolibérale*, Paris, La Découverte, pp. 262-290.

Supiot (A.), 2015, *La gouvernance par les nombres – Cours au Collège de France* 2012-2014, Paris, Fayard.

Thévenot (L.), 1994, « Statistique et politique. La normalité du collectif », *Politix*, n° 7, vol. 25, pp. 5-20.

Vrignaud (P.), 2006, « La mesure de la littéracie dans PISA : la méthodologie est la réponse, mais quelle était la question ? », *Revue française de pédagogie,* n° 157, pp. 27-41.

Ward (M.), 2004, *Quantifying the World. UN Ideas and Statistics*, Bloomington, Indiana University Press.