

Secrétariat technique Permanent
Imm. Kébé Extension 3^{ème} étage
BP 3220, Dakar, Sénégal

Téléphones :
221 33 821 60 22
221 33 822 47 34

Télécopie :
221 33 821 32 26

Courriel :
confemen@confemen.org
site internet :
www.confemen.org

L'évaluation des systèmes éducatifs : pour un meilleur pilotage par les résultats

L'évaluation des systèmes éducatifs : pour un meilleur pilotage par les résultats



Document de réflexion
et d'orientation



**Conférence des ministres de l'Éducation
des pays ayant le français en partage**

L'évaluation des systèmes éducatifs : pour un meilleur pilotage par les résultats



Document de réflexion
et d'orientation

LISTE DES SIGLES ET ACRONYMES	5
Préambule	7
CHAPITRE I : LES ENJEUX DE L'ÉVALUATION	9
1.1. L'ÉVALUATION : UN PROCESSUS COMPLEXE ET MULTIDIMENSIONNEL 11	
1.1.1. Les finalités.....	11
1.1.2. Les domaines d'intervention	13
1.1.3. Les outils	15
1.1.4. Les acteurs et les cibles	16
1.2. L'ÉVALUATION : UNE EXIGENCE POUR LE PILOTAGE DES POLITIQUES PUBLIQUES	19
1.2.1. L'évaluation et les engagements internationaux	19
1.2.2. L'évaluation au service des politiques éducatives	21
1.2.3. L'évaluation et les politiques éducatives au niveau international	25
CHAPITRE II : L'ÉVALUATION DANS L'ESPACE FRANCOPHONE	27
2.1. LES STRUCTURES ET ACTIVITÉS D'ÉVALUATION AU NIVEAU NATIONAL ...29	
2.1.1. Les dispositifs nationaux d'évaluation	29
2.1.2. Les évaluations nationales	32
2.1.3. État des lieux de l'évaluation dans l'espace francophone	35
2.2. LES PROGRAMMES INTERNATIONAUX D'ÉVALUATION	41
2.2.1. Rôles et fonctions des programmes internationaux d'évaluation	41
2.2.2. La coopération entre les programmes internationaux d'évaluation	45
CHAPITRE III : LES ORIENTATIONS	47
3.1. LE MÉMORANDUM SUR L'ÉVALUATION DES SYSTÈMES ÉDUCATIFS	49
3.2. PRINCIPE DIRECTEUR : DÉVELOPPEMENT ET PARTAGE D'UNE CULTURE D'ÉVALUATION	50
3.3. LES PRINCIPES ET OBJECTIFS DES DISPOSITIFS NATIONAUX D'ÉVALUATION ...51	
3.4. LES VOIES D'OPÉRATIONNALISATION	52
3.4.1. L'institutionnalisation de l'évaluation	52
3.4.2. La disponibilité et la prise en compte des informations	53
3.4.3. La mise en place d'une stratégie de communication	54
3.4.4. L'implication et la participation des acteurs aux processus d'évaluation	54
3.4.5. Le réinvestissement des résultats	56
CHAPITRE IV : LES PERSPECTIVES D'ACTION	59
4.1. LES RÉSULTATS ATTENDUS DE L'INSTITUTIONNALISATION DE L'ÉVALUATION ...61	
4.1.1. au plan du pilotage pédagogique	61
4.1.2. au plan du pilotage administratif	61
4.1.3. au plan économique et financier	61
4.1.4. au plan social et politique	62

4.2. DES CONDITIONS POUR L'INSTITUTIONNALISATION DE L'ÉVALUATION	64
4.2.1. Une volonté politique réelle et effective	64
4.2.2. La mise en place de dispositifs d'évaluation	65
4.2.3. La mise à disposition des ressources appropriées	66
4.3. L'ARTICULATION ENTRE LES ACTIVITÉS NATIONALES, RÉGIONALES ET INTERNATIONALES	66
4.4. LE RÔLE DE LA CONFEMEN	68
CONCLUSION	71
BIBLIOGRAPHIE	75

LISTE DES SIGLES ET ACRONYMES

ACDI :	Agence canadienne de développement internationale
ADEA :	Association pour le développement de l'éducation en Afrique
AERES :	Agence d'évaluation de la recherche et de l'enseignement supérieur
AFD :	Agence française de développement
AGEPA :	Amélioration de la gestion de l'éducation dans les pays Africains
APC :	Approche par les compétences
BM :	Banque mondiale
CONFEMEN :	Conférence des ministres de l'Éducation des pays ayant le français en partage
DH-PPTE :	Développement humain – Pays pauvres et très endettés
DEPP :	Direction de l'évaluation, de la prospective et de la performance du ministère de l'éducation nationale
DSRP :	Documents de stratégie pour la réduction de la pauvreté
EPT :	Éducation pour tous
IEA (AIE) :	<i>International Association for the Evaluation of Educational Achievement</i> ou Association internationale pour l'évaluation des acquis scolaires
IIPE :	Institut international de planification de l'Éducation
IMOA-EPT :	Initiative de mise en œuvre accélérée de l'Éducation pour tous
INEADE :	Institut national d'étude et d'action pour le développement de l'Éducation
INES :	<i>International Networks on Education Statistics</i>
INRP :	Institut national de recherche pédagogique
IPN :	Institut pédagogique national
IREDU :	Institut de recherches en économie de l'éducation
LMD :	Licence, master, doctorat
MLA :	<i>Monitoring Learning Achievement</i> ou Suivi permanent des acquis scolaires
OCDE :	Organisation de coopération et de développement économique
OIF :	Organisation internationale de la Francophonie

ONG :	Organisation non gouvernementale
OMD :	Objectifs du Millénaire pour le développement
ONU :	Organisation des Nations unies
PASEC :	Programme d'analyse des systèmes éducatifs de la CONFEMEN
PISA :	Programme international pour le suivi des acquis des élèves
PIB :	Produit intérieur brut
PIRLS :	Programme international de recherche en lecture scolaire
PPTTE :	Pays pauvres très endettés
PREALC :	Programme régional d'emploi pour l'Amérique latine et les Caraïbes
RESEN :	Rapport d'état d'un système éducatif national
SACMEQ :	<i>Southern and Eastern Africa Consortium for Monitoring Educational Quality</i> ou Consortium de l'Afrique australe et orientale pour le pilotage de la qualité de l'éducation
SIGE :	Système d'information pour la gestion de l'éducation
SISED :	Système d'information statistique sur l'éducation
SNRES :	Système national d'évaluation des rendements scolaires
STP :	Secrétariat technique permanent
TIMSS :	<i>Trends in International Mathematics and Science Study</i>
UNESCO :	Organisation des Nations unies pour l'éducation, la science et la culture
UNICEF :	Fonds des Nations unies pour l'enfance
WAEC :	<i>West African Examination Council</i>

Préambule

L'élaboration de ce document s'inscrit dans le cadre de la politique de la Conférence des ministres de l'Éducation des pays ayant le français en partage (CONFEMEN) qui vise à faire connaître ses positions par rapport aux thèmes discutés en session ministérielle à travers la publication d'un document de réflexion et d'orientation.

La thématique de la 51^e session ministérielle tenue à Maurice, en octobre 2004, portait sur l'évaluation des systèmes éducatifs. Elle a permis aux ministres d'échanger sur les enjeux de l'évaluation dans un contexte marqué par les défis de l'Éducation pour tous (EPT). Ainsi, au regard de l'urgence née des engagements pris dans le cadre de l'EPT, de la valorisation des savoir-faire de la CONFEMEN en matière d'évaluation, de l'extension de la réflexion internationale sur l'évaluation à des domaines nouveaux, la CONFEMEN a décidé de s'engager dans la voie d'une réflexion sur une approche élargie des questions d'évaluation ; initiative somme toute légitime et particulièrement opportune.

Les participants à cette session ont constaté avec consensus que l'atteinte des objectifs de l'EPT adoptés au Forum de Dakar en avril 2000, passe par l'adoption de politiques éducatives plus efficaces, c'est-à-dire qui permettent une meilleure transformation des ressources en résultats chez les élèves. Dans cette perspective, l'évaluation a été reconnue comme le principal instrument pour y parvenir : évaluer afin d'identifier les voies par lesquelles les pays peuvent assurer une meilleure transformation des ressources en résultats, évaluer afin de rendre compte à la communauté nationale et internationale des progrès réalisés.

C'est dans ce cadre que les ministres membres de la CONFEMEN se sont engagés — à travers un mémorandum sur l'évaluation des systèmes éducatifs adopté au cours de la 51^e session ministérielle — à favoriser la mise en place de systèmes d'évaluation et à développer une culture de l'évaluation dans les pays membres.

Ce présent document de réflexion et d'orientation fait donc suite à ces échanges. Par sa diffusion, la CONFEMEN entend faire partager sa réflexion quant aux systèmes d'évaluation qu'il convient de mettre en œuvre dans les pays membres tout en affirmant l'engagement de ceux-ci à faire de l'évaluation une réalité dans leurs système éducatifs.

Rappelons que ce document n'a pas pour objectif de fournir une panacée pour la mise en place de systèmes d'évaluation efficaces. Il essaie cependant de mettre l'accent sur des principes essentiels qu'il convient de prendre en considération lors de la mise en place d'un système d'évaluation.

CHAPITRE I :

LES ENJEUX DE L'ÉVALUATION

1.1. L'ÉVALUATION : UN PROCESSUS COMPLEXE ET MULTIDIMENSIONNEL

L'évaluation peut concerner tous les domaines d'un système, et quelle que soit la nature ou la taille du domaine considéré, elle obéit au même processus, et mobilise les mêmes concepts. Elle mesure les écarts entre objectifs et résultats en utilisant pour cela un certain nombre de données (statistiques, rapports divers, rencontres, entretiens de groupe ou individuels, visites de terrain, etc.) qu'elle recoupe de façon à obtenir des réponses fiables à des indicateurs qu'elle s'est donnés.

1.1.1. Les finalités

L'évaluation répond à la double nécessité de rendre compte aux acteurs du système et, plus largement, à l'ensemble de la société et de contribuer à l'amélioration du système éducatif, **précisément à l'amélioration de ses performances, en aidant les décideurs à prendre des mesures adéquates.**

La dimension de «compte-rendu», qui constitue aujourd'hui l'enjeu le plus vif attaché à l'évaluation, implique la mise sur pied d'une véritable politique de communication ou stratégie d'information en direction de tous les acteurs, de tous niveaux, des responsables politiques aux rangs les plus élevés jusqu'aux parents d'élèves, en passant par tous les échelons intermédiaires, au niveau national, régional, local, etc.

De son côté, la fonction de «contribution à l'amélioration des systèmes éducatifs» suppose que l'évaluation puisse fournir une aide efficace, rapidement disponible, à la prise de décisions, au pilotage et à la régulation globale du système. Le Cadre d'action de Dakar parle même de «suivi et d'évaluation continus».

Or, dans beaucoup de cas, on observe que les résultats des évaluations conduites ne sont pas toujours pris en compte, et que leur diffusion laisse beaucoup à désirer, notamment en direction de nombreux acteurs du terrain. C'est largement pour lutter contre cette tendance, fréquemment observée et dénoncée, que l'accent a été mis ces dernières années sur l'évaluation comme outil de **pilotage des systèmes**, et sur l'**évaluation participative**, cherchant à impliquer dans le processus tous les acteurs concernés, des responsables de haut niveau aux parents, en passant par tous les acteurs intermédiaires de l'administration et par ceux de la société civile.

Dans une telle perspective, il est clair que l'évaluation ne saurait donc être considérée comme une activité «périphérique». Elle devrait au contraire être installée au cœur même des systèmes, et y fonctionner comme un outil essentiel des régulations indispensables.

En fait, on convient de plus en plus que dans la chaîne *identification d'objectifs/prise de décision/mise en œuvre et fonctionnement/mesure des résultats obtenus*, des mécanismes de pilotage et de régulation interviennent à chaque étape. L'évaluation, présente elle aussi à chaque étape, constitue l'un des instruments de pilotage les plus fiables.

Les revues sectorielles qui ont été développées à la fin des années 90 peuvent être considérées comme un exemple d'évaluation participative des systèmes éducatifs (voir encadré 1).

Encadré 1 : Les revues sectorielles

Les revues sectorielles en éducation ont pour fonction d'assurer la mise en place et la poursuite d'un véritable partenariat entre le pays et ses partenaires techniques et financiers, fondé sur un diagnostic sectoriel partagé et une vision commune de la politique éducative à conduire sur le court et moyen terme.

Les revues sectorielles constituent un moment d'échanges entre le gouvernement d'un pays et ses partenaires techniques et financiers au sujet du secteur éducatif, en vue de s'assurer que les actions menées ou planifiées sont cohérentes avec les objectifs visés, qui sont réalistes par rapport aux moyens financiers disponibles et compatibles avec la capacité d'absorption du système.

Elles doivent permettre un suivi et un ajustement réguliers de la politique conduite au niveau national, pour une saine articulation entre les engagements politiques pris au niveau national ou international et la mise en œuvre technique à travers le plan annuel d'actions.

Les revues sectorielles doivent permettre de définir clairement les engagements réciproques des partenaires et d'assurer leur suivi. L'exercice des revues sectorielles en éducation :

- s'inscrit dans un processus plus global de suivi-évaluation de la mise en œuvre de la stratégie de réduction de la pauvreté (cohérence des cadrages macro-économique et sectoriel; contribution du secteur de l'éducation à la réduction des inégalités et de la pauvreté) ;

- participe à l'appropriation de la stratégie sectorielle par l'ensemble des acteurs du système éducatif (structures ministérielles centrales et déconcentrées, syndicats d'enseignants, associations de parents d'élèves, etc.).

L'intérêt porté aux revues sectorielles est concomitant aux réformes engagées pour améliorer l'efficacité de l'aide, à savoir :

- le passage d'une conditionnalité *ex ante* de moyens à un partenariat basé sur l'examen *ex post* des résultats ;

- l'inscription des procédures de l'aide dans un processus de coordination et d'harmonisation multi-donateurs ;

- la nécessaire cohérence et articulation des différents instruments de l'aide (aide projet, aide programme, aide budgétaire, etc.).

L'ensemble de ces remarques recoupe largement les stratégies et recommandations faites dans le cadre de nombreuses études ou programmes, qu'il s'agisse du Cadre d'action de Dakar (2000) ou encore des réflexions menées par l'Association pour le développement de l'Éducation en Afrique (ADEA), ou des études des principaux bailleurs de fonds, bilatéraux ou multilatéraux.

Les réflexions en cours prennent de plus en plus en compte ces évolutions dans les approches en évaluation, notamment sur le plan systémique, et s'efforcent de mettre en place les voies et moyens de mieux intégrer l'évaluation au dispositif global de pilotage.

1.1.2. Les domaines d'intervention

Les formes les plus anciennes et les plus «classiques» de l'évaluation concernent **l'évaluation des personnels** (inspections, notations, déroulement de carrière, etc.) et celle des **élèves** (par le biais de notations continues et de la passation d'examens). Les premiers débats sur ces pratiques d'évaluation ont eu lieu au cours des années 70. Selon certains, l'inspection, trop formelle et trop ponctuelle, ne prendrait pas suffisamment en compte les pratiques réelles des enseignants. D'autres, concernant les tests et examens, se sont posés les questions suivantes : qu'évalue-t-on en fait ? Les connaissances et les compétences des élèves ? Ou une copie, produite à un moment donné ?

Le débat sur les limites de l'inspection des personnels et de la notation des élèves a conduit au développement d'autres approches. C'est ainsi que ces dernières années, les approches, en matière d'évaluation, se sont diversifiées et élargies. Il convient donc de les prendre en compte, sans pour autant renoncer aux acquis des approches précédentes.

Ce fut d'abord, l'approche dite qualitative, parfois encore dénommée «**évaluation par les résultats**». Elle est fondée sur les résultats des élèves, niveau par niveau, mesurés au travers d'une batterie de tests spécifiques, et dont l'analyse prend en compte des paramètres variés, comme l'environnement géographique et/ou social, l'âge, le sexe, etc. Cette approche est celle qui a été développée par de nombreux organismes et dans divers programmes, notamment le PASEC.

Il s'agit donc dans cette approche d'évaluer la qualité des apprentissages, mesurée par rapport à des objectifs implicites ou explicites des programmes ou des curricula. Les analyses faites des écarts mesurés entre les objectifs implicites et les résultats constatés permettent de fournir aux décideurs d'utiles éléments propres à faciliter la prise de décision sur des points comme le redoublement, la taille des classes, les types d'équipement, etc., considérés comme autant de variables susceptibles d'intervenir dans la plus ou moins bonne qualité des résultats des élèves, et dont la mise en lumière peut ainsi fournir aux responsables des aides utiles à la décision.

Cependant, cette démarche dite d'analyse qualitative a fait à son tour l'objet d'un débat.

Sans pour autant en méconnaître en aucune façon l'intérêt, certains ont souligné plusieurs points :

- L'évaluation par les résultats vise à dégager des marges de manœuvre, comme l'indique d'ailleurs les auteurs du rapport synthèse PASEC, «on doit certes distinguer deux choses : l'exercice de planification, qui recense globalement besoins et ressources disponibles, et l'exercice effectué par le PASEC, qui recherche à la marge des moyens d'augmenter l'efficacité de l'allocation des ressources» (CONFEMEN, 1999). On notera que dans ce cas de figure, l'exercice de planification lui-même n'est pas objet d'évaluation.
- La démarche d'évaluation par les résultats repose sur la prise d'information auprès de groupes représentatifs, à intervalles et niveaux différents. Dans l'idéal, plusieurs prises d'information seraient nécessaires pour mesurer l'impact effectif de telle ou telle mesure recommandée et mise en œuvre. La

prise d'information, le traitement et l'analyse des données exigent du temps, et les délais sont parfois longs entre le lancement d'une évaluation et la production de ses résultats et recommandations. Ceci risque de limiter le rôle d'aide à la décision qui est aussi celui d'une évaluation.

- Enfin, on fait observer que dans cette approche, si l'on mesure bien l'écart entre les objectifs et les résultats obtenus, les chemins parcourus pour y parvenir (organisation du système, fonctionnement, etc.) constituent une sorte de «boîte noire» dont les composants ne sont pas pris en compte (Pair, 2001).

Pour toutes ces raisons, s'est développée, durant la dernière décennie, une réflexion sur une approche plus «**systémique**» de l'évaluation, et qui vise à englober les domaines suivants :

- **les élèves et les étudiants**, en examinant les formes les plus classiques de l'évaluation, mais aussi en tenant le plus grand compte des apports de l'évaluation qualitative des résultats ;
- **les personnels**, en prenant en compte leur degré de formation initiale et continue, les pratiques pédagogiques, les capacités d'innovation, les incitations possibles, etc. ;
- **les structures et leur mode de gouvernance**, sur les plans national, régional ou local, celui des écoles et des établissements ;
- **le système éducatif national**, dans divers aspects de son fonctionnement, comme les expériences, les innovations, les réformes mises en œuvre, les choix de politique éducative, etc.

L'évaluation «systémique» peut s'appliquer à l'ensemble d'un système éducatif ou d'une politique éducative, à l'une de leurs parties, géographique ou sectorielle ou à des unités comme l'établissement scolaire, sous réserve de ne pas oublier les interactions entre les différents niveaux et les différents domaines étudiés.

Quelle que soit sa forme et quel que soit le domaine étudié, l'évaluation obéit à la même **méthodologie** :

- construction d'un «réfèrent» — objectifs, critères, normes — et d'un «référé» — résultats, situations observables à un moment donné de la mise en œuvre ;
- mesure des écarts entre réfèrent et référé, en s'appuyant sur des «données» (statistiques, rapports, entretiens, etc.) et au moyen «d'indicateurs» (pertinence, efficience, efficacité, cohérence, effectivité, impact, visibilité, durabilité, etc.) ;

et remplit, tel qu'il a été mentionné précédemment, **une double fonction**, d'une part, de rendre compte aux acteurs du système, et plus largement, à l'ensemble de la société et, d'autre part, de contribuer à l'amélioration du système éducatif, ce qui implique qu'elle ne

peut plus être considérée comme une activité «périphérique».

L'encadré suivant présente les facteurs qui peuvent déterminer le choix d'un type d'évaluation :

Encadré 2 : Choix d'un type d'évaluation	
FACTEURS DÉTERMINANTS	TYPES D'ÉVALUATION
Agent	■ interne (ou auto-évaluation) ■ externe (ou évaluation indépendante)
Moment	■ <i>ex-anté (ou appréciation)</i> ■ mi-parcours (intermédiaire) ■ finale (évaluation menée à l'achèvement) ■ <i>a posteriori (évaluation menée après l'achèvement)</i>
Champ d'application	■ de projet ■ sectorielle ■ thématique ■ de programme ■ de politiques / stratégique ■ de processus

1.1.3. Les outils

Pour conduire une évaluation, il faut disposer d'information. Cette dernière doit être collectée au moyen d'un certain nombre d'outils qui sont essentiellement :

- **l'utilisation de la documentation** (travaux antérieurs, statistiques officielles, études existantes, etc.) ;
- **les observations directes** (sur le terrain) ;
- **l'analyse des indicateurs** (de type «tableau de bord» ou indicateurs simples et pertinents choisis pour le suivi des actions et l'appréciation de leurs effets) ;
- **les enquêtes et entretiens** (pour recueillir des informations complémentaires, faire le point à un moment donné ou mesurer des évolutions) ;
- **les avis des bénéficiaires** (à travers des réunions ou entretiens groupés ou séparés pour permettre à certains groupes de mieux s'exprimer).

Dans l'optique de la scolarisation universelle, l'amélioration de la qualité des services éducatifs constitue une urgente nécessité. Il faut donc pouvoir évaluer la manière dont les services éducatifs sont offerts d'une part, et apprécier le niveau d'acquisition des élèves d'autre part. D'où la nécessité de mettre en place des instruments d'évaluation ciblés sur les résultats pour s'assurer que des progrès dans la transformation des ressources en résultats scolaires sont bien réalisés. Les approches pour juger des résultats de l'apprentissage varient. Certaines évaluent les résultats de l'apprentissage au moyen d'examens destinés à déterminer la maîtrise des matières inscrites au programme. D'autres évaluent la maîtrise d'une série de niveaux d'apprentissage minimaux séquencés ou de compétences d'apprentissage essentielles. D'autres encore ont recours à des tests normalisés de

compétences de base ou à des tests d’alphabétisme fonctionnel d’arithmétique et de «savoir-faire utiles». Les **examens de fin de cycle** et les **tests standardisés**, bien que ne consistant pas en une mesure parfaite des connaissances ou des acquis des élèves, peuvent constituer des indicateurs très utiles si on les emploie avec précaution.

Le choix des outils, grâce auxquels sont menées les évaluations, n’est pas neutre. En effet, chaque dispositif est construit autour d’une finalité particulière, et son ancrage institutionnel est susceptible d’influer sur le type de résultats obtenus et sur leur exploitation. L’outil est au cœur des problématiques d’évaluation.

La CONFEMEN a réalisé une enquête auprès des pays membres à propos des systèmes d’évaluation mis en place (ou non) dans les secteurs éducatifs.

Les réponses à ce questionnaire¹ montrent que les outils employés pour l’évaluation des résultats sont principalement des tests (sauf pour les évaluations réalisées par les corps d’inspection sur la base d’observations et d’entretiens), soit réalisés à une échelle nationale (cas du dispositif «système national d’évaluation des rendements scolaires» (SNERS) au Sénégal), soit des outils standardisés relevant d’enquêtes internationales (PASEC et Suivi permanent des acquis scolaires : *Monitoring Learning Achievement ?* MLA). **Dans tous les cas, il s’agit d’outils dont l’objet est la mesure de performances des élèves, et non la mesure de performances des établissements.** Toutefois, comme c’est le cas pour le PASEC, l’évaluation peut déboucher sur des modèles d’établissements peu coûteux et efficaces.

On note également que l’exploitation des résultats des examens à des fins de pilotage ou de décision n’est jamais inscrite dans le paysage des politiques nationales d’évaluation interrogées : les résultats d’examen font partie des champs d’évaluation cités par 15 États et/ou gouvernements sur les 23 répondants, mais sans rétroaction sur les politiques ou les décisions locales, à la différence de ce qui est argué pour les tests nationaux ou internationaux.

1.1.4. Les acteurs et les cibles

1.1.4.1. Les acteurs de l’évaluation

Entre l’évaluation qui serait totalement pilotée par le bailleur de fonds sans implication d’autres acteurs (évaluation externe au sens plein du terme) et l’auto-évaluation réalisée par les bénéficiaires eux-mêmes, les formules mixtes sont nombreuses. Selon l’enquête préparatoire menée dans le cadre de ce travail, dans la grande majorité des cas, les structures préexistantes ont été chargées des opérations périodiques d’évaluation du système au niveau national. Ce sont principalement les corps d’inspection (neuf cas répertoriés sur les 23 répondants) ou les services de planification (huit cas répertoriés sur les 23 répondants). Parfois, des cellules créées à la faveur de la passation de tests PASEC demeurent comme le principal acteur de l’évaluation (par exemple, à Madagascar, au Tchad et au Togo). Des agences spécifiquement créées, en charge d’une approche globale de la qualité, et sans lien hiérarchique technique avec les structures opérationnelles du ministère, existent ponctuellement : c’est le cas au Luxembourg et en Roumanie.

¹ Notons que 23 États et gouvernements membres de la CONFEMEN ont répondu au questionnaire sur l’état des lieux des dispositifs d’évaluation des systèmes éducatifs nationaux.

Dans tous les cas, la participation à des tests internationaux² mobilise fortement les acteurs nationaux, qui participent à l'ensemble des opérations, de la conception ou de l'adaptation des séries d'items à leur administration et à l'exploitation des résultats. Pour autant, cette participation n'est pas forcément inscrite dans une périodicité régulière et récurrente, et ne génère pas toujours l'émergence d'équipes stables de professionnels qui développeraient des outils nationaux ou locaux.

Au niveau régional, la plupart des pays répondant à l'enquête ne sont pas dotés de structures permanentes dédiées à l'évaluation. La fonction n'est assurée au niveau régional que dans 13 cas sur 23, elle l'est généralement par des structures à vocation plus générale (notamment direction régionale, antennes des directions de la planification et inspection régionale).

Au niveau local, cette fonction n'est plus représentée que dans huit États et/ou gouvernements sur les 23 répondants, dans les mêmes conditions que pour l'échelon régional.

Assurément, l'enjeu principal du développement des pratiques d'évaluation réside dans une plus grande participation des acteurs aux niveaux local et régional, ce qui suppose une structuration renouvelée des services concernés.

Dans presque tous les cas, les pays ayant répondu à l'enquête soulignent l'**insuffisance des ressources humaines** (18 réponses sur 23 en ce sens), cette insuffisance étant aussi bien en termes d'effectifs qu'en termes de formations des acteurs. Au titre des souhaits et suggestions pour l'amélioration des dispositifs existants, la mention des renforcements de capacités par le truchement de formations spécifiques vient en tête des préoccupations.

1.1.4.2. Les cibles de l'évaluation

❖ Les destinataires

On peut distinguer plusieurs catégories de destinataires auxquels l'évaluation permettrait de «rendre compte» :

- une première catégorie de destinataires est constituée par les **financeurs** :
 - ✓ pouvoirs publics nationaux, États, collectivités locales et instances de contrôle relevant d'eux ;
 - ✓ bailleurs de fonds étrangers dans le cadre d'une coopération bi ou multilatérale, ou de fonds privés étrangers ou nationaux ;
- une seconde catégorie est composée des **acteurs de l'action** évaluée, en l'occurrence, dans le cas de l'éducation, les enseignants, les administrateurs aux différents niveaux de la hiérarchie des responsabilités, du directeur d'école au ministre, ou les différentes instances de contrôle interne du système éducatif ;

² Parmi les pays ayant répondu à l'enquête, seuls Djibouti, la Guinée-Bissau et la Côte d'Ivoire n'ont participé à aucune évaluation internationale depuis 2000.

- une troisième catégorie est celle des «**usagers**» : les familles, les élèves eux-mêmes, les communautés, les entreprises (qui utilisent les «produits» de l'école), etc.

❖ Les objets de l'évaluation

Ce que l'on évalue étant toujours les résultats d'une action, ce sont nécessairement aussi des acteurs (des acteurs, non des personnes) que l'on évalue, quelle que soit leur position dans le système. C'est pourquoi l'évaluation ne peut aboutir à des améliorations que si les acteurs ne sont pas seulement objet mais partie prenante à l'évaluation, non pas seulement à sa réalisation, mais aussi à sa conception. D'où l'idée d'une **évaluation formative et participative** dont le but est d'assurer la progression de l'apprenant dans une démarche d'apprentissage. Parmi les applications possibles du principe de l'évaluation formative se trouve la mise en place de relances pédagogiques concrètes à partir d'une évaluation diagnostique conduite par les maîtres eux-mêmes. Dans ce principe, il s'agit surtout de veiller à l'acquisition réelle de connaissances et de compétences.

Quand on se propose d'évaluer la qualité d'un système éducatif, l'objet premier est ce qu'on appelle les **performances scolaires**. Les méthodes pour les établir sont rigoureuses, dès lors qu'elles peuvent s'appuyer sur des données fiables. Les méthodes de mesure de cette fiabilité sont également rigoureuses et permettent par là même de mesurer la fiabilité des résultats constatés. Mais évaluer n'est pas seulement constater, mesurer des écarts, calculer des ratios, des coûts unitaires, c'est aussi interpréter et expliquer. L'évaluation a donc une fonction de **diagnostic** qui permet d'avancer des hypothèses explicatives en vue de remédiations.

Ces hypothèses conduisent à définir d'autres objets d'évaluation, en vue de déterminer les points forts et les points faibles de tel ou tel aspect de la politique éducative ; par exemple, la politique de formation des maîtres, le rôle des manuels scolaires, etc. C'est ce qu'on appelle des évaluations **thématiques**. En effet, les moyennes nationales recouvrant généralement de très grandes disparités géographiques, l'évaluation doit mesurer et expliquer ces disparités. Les évaluations régionales ou locales impliquent des analyses complexes — économiques, sociales, culturelles, etc. — relevant d'approches pluridisciplinaires. En outre, les évaluations thématiques et les évaluations régionales conduisent à définir comme objets à évaluer **l'organisation, le fonctionnement et la gestion du système éducatif** à ses différents niveaux, de l'école à l'administration centrale et aux instances en charge de définir, et non plus seulement d'appliquer, la politique éducative. On comprend dès lors la difficulté d'une évaluation qui soumet à examen critique l'exercice du pouvoir politique lui-même. Cette difficulté, mais aussi, sa possible efficacité si le pouvoir politique accepte de donner l'exemple, peut créer une «spirale vertueuse» entraînant les acteurs de terrain.

Ce serait là non pas la condition *sine qua non*, du moins un puissant facteur d'accélération, pour que **l'évaluation des établissements** devienne réellement la stratégie efficace qu'on attend pour une amélioration globale du système.

1.2. L'ÉVALUATION : UNE EXIGENCE POUR LE PILOTAGE DES POLITIQUES PUBLIQUES

L'éducation ne fait pas figure de secteur d'exception au sein des interventions politiques ou sociales des États. L'évaluation des politiques publiques, ainsi que le développement de cadres budgétaires rénovés destinés à faciliter l'évaluation, apparaît depuis un quart de siècle comme une exigence qui accompagne, voire encadre, la définition de l'action des États.

C'est donc l'ensemble des politiques publiques qui est aujourd'hui régulièrement questionné au regard des résultats atteints et de leur relation avec les moyens employés.

1.2.1. L'évaluation et les engagements internationaux

Le Cadre d'action de Dakar (2000), les objectifs du Millénaire pour le développement (2000) ainsi que l'initiative de l'allègement de la dette des pays pauvres très endettés (l'initiative PPTE) fournissent des objectifs généraux et des orientations stratégiques globales quant à l'orientation et la mise en action de l'évaluation. Cependant, les spécificités des pays et la diversité de leurs situations nécessitent l'élaboration de stratégies différenciées qui font l'objet d'un suivi et d'une évaluation continue.

Il convient de rappeler que 22 pays du monde en développement ayant satisfait, par le caractère soutenable des politiques publiques qu'ils mettent en œuvre, aux critères d'éligibilité à l'initiative PPTE³, verront leur dette extérieure réduite en moyenne de 45%, ce qui, malgré des disparités considérables entre pays, permet de dégager des ressources importantes au niveau des pays (de l'ordre de 12 % à 26 % des recettes fiscales en moyenne). Cette opportunité crée pour le secteur éducatif un enjeu double : récupérer des ressources de la réduction de la dette, mais surtout faire la preuve qu'elles seront utilisées de manière efficiente et équitable. On retrouve ici la nécessité de mettre en œuvre des politiques et des stratégies appropriées. L'idée centrale étant qu'il s'agit, pour ces pays, de considérer les ressources additionnelles rendues disponibles par l'initiative PPTE comme une opportunité de se pencher qualitativement sur le cœur de leurs systèmes.

Dans le secteur de l'éducation, l'évaluation désigne un mécanisme qui :

- a pour finalité de mettre en relation la performance globale des systèmes éducatifs avec les moyens mobilisés, afin de dégager, de façon empirique et au moyen de comparaisons, des marges de manœuvre et d'amélioration pour les systèmes ;
- emploie des méthodes qualitatives et quantitatives permettant de comparer les résultats dans le temps et dans l'espace, et d'isoler des variables utiles à la détermination des politiques ;
- s'intègre dans un processus plus large de gestion par les résultats associant acteurs et bénéficiaires du système dans la recherche continue de voies d'amélioration.

³ Il s'agit de pays ayant atteint le point de décision, c'est-à-dire de pays disposant d'une stratégie de réduction de la pauvreté et d'un programme d'éducation sectorielle, adoptés par le gouvernement et endossés par les bailleurs de fonds conformément au cadre indicatif.

Les pays engagés dans le processus d'initiative d'allègement de la dette des PPTE ou l'Initiative de mise en œuvre accélérée de l'Éducation pour tous (IMOA-EPT) ont le plus souvent élaboré et voté une loi d'orientation ou ont adopté, en Conseil des ministres, une lettre de politique sectorielle. Dans ce cadre, ils ont établi des plans ou programmes à long terme, le plus souvent décennaux.

L'ensemble de ces documents (Documents de stratégie de réduction de la pauvreté (DSRP), lois d'orientation, programmes décennaux, etc.) présente un double avantage :

- Ces documents manifestent l'affirmation d'une politique nationale, fournissant des objectifs, un cadre d'action, un horizon à moyen et long terme et des étapes intermédiaires à atteindre. Ceci clarifie nettement le paysage éducatif tout en étant de nature à inciter les bailleurs de fonds à s'impliquer davantage dans le processus.
- Ils permettent de répondre aux critiques souvent formulées à l'encontre de «l'approche projets» (taille restreinte des projets, chevauchements, dispersion, faible impact, etc.). Même si le volume global de l'aide-projet en provenance de la coopération internationale n'intervient qu'à la marge dans le secteur, l'élaboration de stratégies claires, traduites dans les programmes décennaux du pays permet une meilleure coordination de l'aide et peut ainsi lui assurer un impact plus important.

Les décideurs se sont engagés dans un dialogue avec des bailleurs de fonds et sont ainsi amenés à rendre compte de leur engagement, tout comme ils le font à l'égard d'organismes internationaux qui assurent le suivi des politiques auxquelles lesdits bailleurs participent également. Ceci vaut, par exemple, pour l'élaboration des DSRP ou des programmes décennaux de l'éducation, qui sont d'ailleurs soumis à des examens périodiques.

Mais dans le même temps aussi, les États se doivent d'évaluer leurs propres politiques non seulement parce que des engagements extérieurs ont été pris, mais aussi parce que, au travers des lois d'orientation du système éducatif, des engagements ont été pris devant la représentation nationale et, plus largement, devant la société civile⁴. Les États doivent ainsi se doter d'un dispositif national d'évaluation, afin de piloter leur action en s'appuyant sur un «tableau de bord» fait d'une série d'indicateurs. Ceux-ci, comme tous les indicateurs, doivent être fiables, pertinents, objectivement vérifiables et sensibles, c'est-à-dire capables de mesurer dans le temps les variations d'une situation donnée.

Il est donc clair, qu'ici encore, l'évaluation joue un rôle central : des engagements publics ayant été pris, il importe que les résultats attendus soient atteints. Un pilotage s'appuyant sur un dispositif permanent d'évaluation est l'un des moyens de réguler l'action, d'apporter de nécessaires rectifications et de se donner de meilleures chances d'atteindre les objectifs fixés.

⁴ On entend par «société civile» la représentation ordonnée et structurée de l'ensemble global de la société qui se distingue de la «communauté» au sens large. La société civile s'exprime ainsi à travers des associations, des ONG, des syndicats ou des groupes représentatifs. La société civile peut être en demande pour l'évaluation des services publics éducatifs, elle peut y être associée, ayant des points de vue particuliers à faire valoir. A défaut, les politiques mises en place par les États pourraient être perçues comme éloignées des besoins réels des populations, et n'obtenir que des résultats décevants.

Cette activité d'évaluation se distingue de l'exercice global de planification en ceci qu'elle implique un jugement de valeur sur les objectifs eux-mêmes et sur la mise en œuvre. Par exemple, si l'on se fixe comme objectif de passer d'un taux de scolarisation des filles de 40 % à 70 % en cinq ans, plusieurs questions se posent, en termes d'évaluation, comme :

- S'agit-il d'une moyenne nationale ?
- A-t-on pris la mesure de disparités régionales ?
- Les premiers exercices montrent-ils des progressions telles que l'on peut espérer atteindre l'objectif dans l'agenda fixé ?
- Sinon, quels obstacles sont apparus ?
- Comment y remédier pour rester dans les temps ?
- *In fine*, l'objectif et le calendrier arrêtés étaient-ils bien réalistes ?

Qu'il s'agisse de planification globale ou de mesure des performances des élèves, un processus permanent d'évaluation constitue un instrument de pilotage indispensable à la poursuite d'objectifs et d'engagements ambitieux, de par leur ampleur et de par leur calendrier.

1.2.2. L'évaluation au service des politiques éducatives

Dans le contexte particulier de la gestion des systèmes éducatifs, l'évaluation a souvent cherché à mesurer le degré d'atteinte des missions confiées à l'école afin de servir à leur pilotage. À cet égard, force est de constater que les mutations sociales et économiques redéfinissent continuellement les missions du système éducatif. Cependant, il existe un consensus visant à reconnaître que parmi les missions fondamentales d'un système éducatif, il y a :

- l'accès au service éducatif ;
- l'équité dans l'offre de service éducatif ;
- la qualité des apprentissages des élèves.

Dans le Cadre d'action de Dakar (2000), la communauté internationale reconnaît la légitimité de ces axes en engageant les pays à la recherche d'une éducation de qualité pour tous. Elle les a donc traduits en objectifs chiffrés et mesurables à atteindre pour 2015, définissant ainsi divers indicateurs pour l'accès, l'équité et la qualité des apprentissages.

1.2.2.1. L'évaluation et la qualité

Une action est jugée de qualité lorsqu'elle satisfait à la fois aux trois critères : efficacité, efficience et pertinence.

L'efficacité interne et externe

L'efficacité fait référence à la réussite, au succès. Elle mesure le degré d'atteinte et de réalisation des objectifs ; d'où l'importance d'avoir des objectifs clairs et précis dès le départ. L'intérêt est de mesurer des écarts et de pouvoir les analyser. À cet effet, il convient également de ne pas négliger le point de vue des bénéficiaires. Il s'agit en l'occurrence de ceux qui ont participé activement à l'action, ceux qui ont partiellement participé, ceux qui refusent de participer ou qui ont été exclus, etc. Pour le système éducatif, l'efficacité se situe à deux niveaux : l'un interne et l'autre externe.

L'efficacité interne d'un système éducatif, dans le contexte des missions confiées à l'école, s'intéresse à la mesure des relations entre les résultats obtenus et les ressources engagées. Elle

s'apprécie donc en termes qualitatifs et quantitatifs. Sur le plan qualitatif, l'efficacité interne mesure le rapport entre les apprentissages scolaires acquis et les objectifs pédagogiques assignés au système. Sur le plan quantitatif, elle mesure la capacité du système à maintenir en son sein, et pour la durée du cycle prévu, tous les enfants qui y sont inscrits. C'est d'ailleurs cet aspect de l'efficacité interne qui est souvent mis en avant. Ainsi, moins il y a de redoublement et d'abandons, plus le système est jugé efficace. À l'inverse, des taux d'abandons et de redoublements élevés reflètent une inefficacité interne et peuvent souvent traduire une iniquité de la qualité de l'encadrement pédagogique des groupes d'élèves en difficultés scolaires. Ils traduisent, dans ce cas, des coûts unitaires élevés de scolarisation.

L'efficacité externe, quant à elle, fait référence à l'adéquation entre le système éducatif et le monde du travail. Elle sert à mesurer la capacité du système éducatif à former et à préparer les élèves à l'insertion dans la vie sociale et dans la vie économique. À ce propos, dans bon nombre de pays, le constat est préoccupant. Les «produits» de l'école ne disposent pas toujours de compétences réelles leur permettant de répondre aux besoins de la société et de l'économie. Il est nécessaire que ces «produits» soient en mesure de s'adapter au développement socio-économique avec ses multiples mutations.

Dans les pays en développement, la faiblesse dans l'efficacité (interne comme externe) a généralement pour conséquence de détourner les populations de l'école, un retrait de confiance qui est surtout manifeste en milieu rural. Par conséquent, l'amélioration de l'efficacité des systèmes éducatifs s'avère être d'une impérieuse nécessité.

L'efficience

L'efficience interroge sur la relation entre les moyens et les avantages tirés. Autrement dit, dans quelle mesure les ressources (financières, techniques, organisationnelles et humaines) ont-elles été bien mobilisées en temps voulu et converties en réalisations au moindre coût ? Les résultats ont-ils été à la mesure des sommes dépensées. Bref, l'efficience est l'appréciation du résultat et de la démarche choisie par rapport aux ressources mobilisées.

La pertinence

La pertinence examine le bien fondé de l'action au regard des objectifs et enjeux déterminés au départ. Elle permet de choisir la solution idoine au problème posé. L'analyse de la pertinence consiste donc à préciser jusqu'à quel point les objectifs de l'action de développement correspondent aux besoins du pays, aux priorités et politiques de développement. Toutefois, les trois dimensions complémentaires que sont l'efficacité, l'efficience et la pertinence, sont souvent inégalement mises en relief.

Accorder plus d'importance à l'efficience comporte le risque de ne prendre en compte que des critères liés à la mobilisation des ressources. S'intéresser uniquement à l'efficacité sans prendre en compte les coûts risque d'être ruineux pour l'ensemble de la société. Par contre, la pertinence permet de juger l'ensemble des effets et non seulement l'écart entre les résultats et les objectifs visés. Elle conditionne ainsi l'efficience et l'efficacité.

Dans le cadre d'une évaluation systémique, les trois dimensions, *efficacité*, *efficience* et *pertinence* doivent être appréciées de manière conjointe afin de fournir une base solide aux recommandations et enseignements tirés de l'expérience et d'améliorer la qualité de l'action

éducative notamment. Ainsi, dans l'exercice d'évaluation, il importe de mettre l'accent sur ces trois dimensions, et ce, depuis la planification des activités d'évaluation jusqu'à leur exécution, depuis la conception des instruments de collecte des données jusqu'à la diffusion de rétro-information à des fins d'apprentissage.

C'est dire que toute évaluation relève d'une décision qui n'est d'abord ni technique ni comptable, mais politique, car elle renvoie implicitement ou explicitement aux finalités de l'action politique en général et à ses priorités réelles. En l'occurrence : quelle est la priorité réellement accordée à l'éducation ? Comment se justifie le rang qu'on lui accorde par rapport aux autres politiques qui répondent aux autres besoins de la société ?

1.2.2.2. L'évaluation, instrument de pilotage des systèmes éducatifs

Comme souligné dans les finalités de l'évaluation, abordées plus haut, un objectif fondamental du processus d'évaluation est de déboucher sur la prise de décision. En effet, à Jomtien en 1990, puis à Dakar en 2000, les États et gouvernements se sont engagés à garantir l'atteinte des résultats mesurables de l'EPT en plaçant l'évaluation au cœur de tout processus de décision.

Dans ce rôle de levier à la prise de décision, les résultats d'une évaluation contribuent au pilotage des systèmes éducatifs en identifiant les politiques et stratégies éducatives les plus efficaces dans un contexte de rareté des ressources. Il s'agit donc en général d'identifier les modèles d'écoles peu coûteux qui garantissent une éducation de qualité pour tous.

À cet égard, les évaluations menées par le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) rejoignent d'autres travaux de la littérature empirique et mettent en lumière un ensemble de facteurs récurrents dans l'amélioration de la qualité des apprentissages des élèves. Il s'agit par exemple :

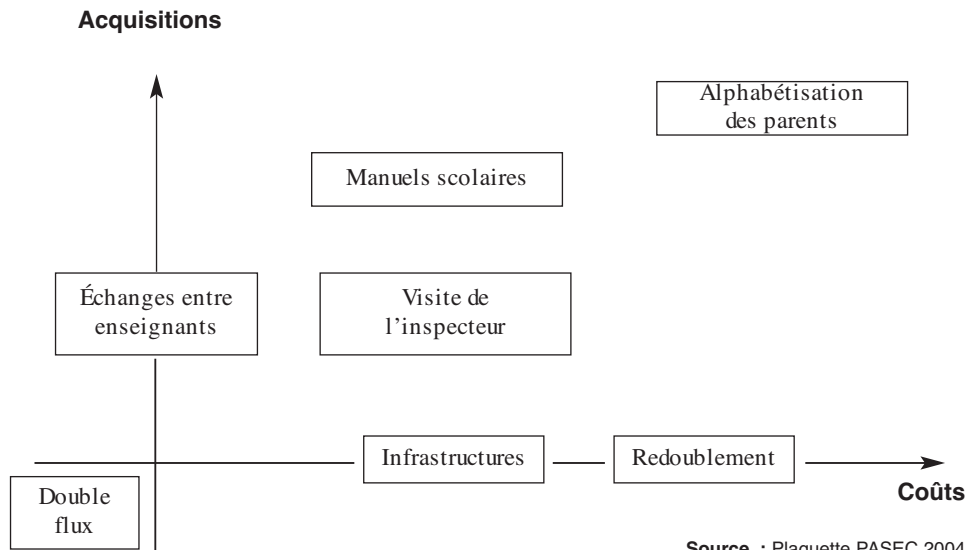
- des politiques de dotation en manuels scolaires qui ont un effet positif sur la qualité des acquisitions scolaires des élèves lorsqu'une stratégie de distribution et d'utilisation par les élèves et enseignants est assurée ;
- des échanges de bonnes pratiques entre enseignants dans le cadre des réunions pédagogiques organisées au sein de l'école ;
- des visites de l'inspecteur aux enseignants et au directeur ;
- de l'alphabétisation des parents des élèves ;
- des politiques de recrutement, de formation, de suivi et de motivation des enseignants.

Cependant, ces indicateurs ne suffisent pas pour piloter efficacement un système éducatif de façon homogène d'un pays à l'autre. En effet, un enseignement majeur tiré des évaluations est celui du poids des spécificités nationales et l'importance des interactions qui lient les facteurs entre eux. De ce fait, certains facteurs peuvent s'avérer déterminants pour la qualité des apprentissages scolaires dans un État alors qu'ils le sont moins dans un autre. Chaque État devrait donc disposer d'au moins une évaluation diagnostique lui permettant d'identifier les facteurs qui influent sur la qualité des apprentissages scolaires de ses élèves afin d'adapter les politiques à son contexte.

Aussi, en raison des interactions entre les facteurs, on ne peut recommander le pilotage des systèmes en considérant les facteurs indépendamment les uns des autres. Le caractère

multidimensionnel de la qualité de l'éducation oblige à conduire une politique qui prend en compte tous les facteurs simultanément. En reprenant, les résultats d'une quinzaine d'études menées dans des pays d'Afrique subsaharienne de la CONFEMEN et, sous l'arbitrage des coûts liés à la traduction en politiques de ces résultats, le PASEC illustre de façon approximative le rapport coût efficacité de quelques facteurs clés à travers le graphique suivant.

Graphique n° 1 : Rapport coût efficacité de quelques facteurs à la lumière de certaines études PASEC



Source : Plaquette PASEC 2004

Ce graphique met en évidence le fait qu'il est possible, grâce à certains choix de politiques éducatives, d'améliorer la qualité de l'éducation sans que cela n'entraîne des coûts démesurés. Il existe donc des marges de manœuvre pour le pilotage des politiques éducatives et on ne saurait trop insister sur l'importance de l'évaluation pour les identifier dans le contexte de chaque État et gouvernement membre.

Enfin, pour servir efficacement au pilotage des politiques éducatives, l'évaluation doit être diverse et variée. En effet, on ne peut pas couvrir avec la même précision tous les aspects du système éducatif au moyen d'un seul type d'évaluation. C'est notamment le cas lors de la mesure et de l'interprétation de l'impact de l'enseignant sur la qualité des apprentissages scolaires des élèves. Ceci s'explique par le fait que les méthodes quantitatives peuvent difficilement cerner, au moyen d'un questionnaire, l'ambiance pédagogique qu'un enseignant instaure dans sa classe. Ainsi, cette évaluation quantitative ne pourra pas servir pour identifier les aspects régissant l'animation pédagogique d'une classe qui font la différence en termes de qualité des apprentissages des élèves. C'est ce qu'on appelle communément «l'effet maître» dans la qualité des apprentissages, au-delà des dimensions de formation initiale et professionnelle ou d'ancienneté qui sont aisément mesurables. Il conviendrait donc

de rapprocher «l'effet maître» avec les pratiques d'enseignement et les caractéristiques personnelles des maîtres (personnalité, comportement, engagement, etc.) pour fonder de nouvelles procédures de recrutement, de formation et d'encadrement des futurs enseignants. Une telle action relève d'une extension des programmes d'évaluation des déterminants de la qualité (de type PASEC ou SACMEQ).

1.2.3. L'évaluation et les politiques éducatives au niveau international

Les enjeux liés à l'évaluation au niveau international sont principalement l'information et le classement qu'on peut faire des pays. Cet aspect des évaluations internationales oblige ces dernières à obéir à un certain nombre de critères devant garantir la comparabilité des résultats d'un pays à l'autre. En effet, pour opérer des comparaisons au niveau international, les populations cibles, la procédure d'échantillonnage, les tests, les consignes de corrections et les procédures de traitement et d'analyse des données doivent être documentées et standardisées (identiques) dans tous les pays.

À l'aune des années 90, et sous l'impulsion des recommandations de Jomtien, plusieurs programmes standardisés d'évaluations nationales des acquis des élèves ont été mis en place. On peut principalement citer :

- le programme de Suivi permanent des acquis scolaires des élèves ou *Monitoring Learning Achievement* (MLA) mis en place par l'UNESCO ;
- le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) ;
- le Consortium de l'Afrique australe et orientale pour le pilotage de la qualité de l'éducation ou *Southern and Eastern Africa consortium for monitoring educational quality* (SACMEQ).

Pour chacun de ces programmes, une méthodologie standardisée a été appliquée dans les pays évalués. Des comparaisons sont donc possibles entre les résultats moyens des élèves au niveau international. La richesse de la comparaison internationale, lorsque les conditions qui président à sa validité sont respectées, vient du fait que les analyses montrent comment différentes politiques éducatives peuvent conduire à des résultats différents. Ceci peut enrichir les marges de pilotage des planificateurs et décideurs politiques, mais tout en tenant compte du contexte de leur pays.

Aussi, pour ces programmes, les pays couverts n'ont souvent pas participé à l'évaluation à la même période. Cette donnée implique que le planificateur ou l'acteur soit souvent obligé de comparer les résultats de pays qui n'ont pas été évalués à la même année. Ce qui rend d'autant plus délicat les conclusions et enseignements tirés de la comparaison internationale.

En plus des enquêtes de type national dont la méthodologie est standardisée à des fins de comparaison comme c'est le cas pour le MLA, le SACMEQ et le PASEC, il existe une seconde catégorie d'enquêtes internationales menées par des programmes, tels que PIRLS et TIMMS de l'IEA, et PISA de l'OCDE. La principale caractéristique de ces dernières réside dans leur programmation. En effet, ces enquêtes sont programmées pour se tenir dans plusieurs pays simultanément suivant une méthodologie et des instruments standardisés. En 2000, trente-deux pays ont participé aux enquêtes PIRLS et TIMMS de l'*International Association for the Evaluation of Educational Achievement* (IEA) connue en français sous le nom d'Association internationale pour l'évaluation des acquis scolaires (AIE). En 2003, ils

étaient au nombre de 43 pour un total de 275 000 élèves (CIEP, 2006).

L'intérêt des médias pour les comparaisons internationales, dû en partie aux critiques que ces dernières soulèvent a amené les programmes internationaux à redéfinir avec plus de précision ce qu'ils évaluent et à diffuser la méthodologie des études ainsi que les précautions à prendre dans l'interprétation des résultats. Ainsi, pour être plus crédibles, compréhensibles et mieux acceptés, les programmes veillent de plus en plus à maintenir à jour un référentiel de compétences ou savoirs en conformité avec les profils de sortie des élèves définis par chaque pays.

À cette date, peu de pays africains participent aux évaluations internationales à grande échelle comme celles de l'OCDE et de l'AIE. Cet état de fait peut s'expliquer par la différence des réalités socio-économiques et culturelles entre la majorité des pays du Sud et ceux du Nord. On peut dès lors s'interroger sur l'opportunité et les objectifs d'évaluations à grande échelle pour des pays du Sud.

Indépendamment de leurs différences méthodologiques et mode de fonctionnement, les programmes internationaux de l'AIE, de l'OCDE, du PASEC, du MLA, du SACMEQ, etc., cherchent à produire à l'intention des décideurs, chercheurs et autres acteurs, des indicateurs internationaux sur la qualité de l'enseignement pour une régulation des systèmes éducatifs. En effet, dans le contexte de la globalisation, «le *benchmarking* apparaît logiquement incontournable» (CIEP, 2006).

Cependant, au-delà des indicateurs de la qualité des acquisitions scolaires, on distingue, suivant la classification du BRED 2005, deux autres grandes familles d'indicateurs qui complètent l'information sur l'efficacité des systèmes éducatifs. Il s'agit des indicateurs de rétention et d'attraction. Les indicateurs de rétention font référence à la capacité du système à maintenir les élèves scolarisés tout au long de leurs apprentissages avec le minimum de redoublement. Les indicateurs d'attraction, quant à eux, décrivent la capacité d'un établissement à scolariser la population des enfants de sa zone d'attraction.

CHAPITRE II :

L'ÉVALUATION DANS L'ESPACE FRANCOPHONE

Afin de formuler des recommandations en faveur de l'évaluation pour tous les pays de l'espace francophone, il est nécessaire d'identifier des caractéristiques communes aux activités et structures d'évaluation menées dans cet espace, tant au niveau national qu'international.

2.1. LES STRUCTURES ET ACTIVITÉS D'ÉVALUATION AU NIVEAU NATIONAL

L'état des lieux non exhaustif que la CONFEMEN a réalisé permet, d'une part, de constater des disparités sur le plan des dispositifs nationaux chargés de l'évaluation et, d'autre part, de mettre en relief les défis qui interpellent les États et gouvernements membres dans ce domaine.

2.1.1. Les dispositifs nationaux d'évaluation

L'analyse montre que si des pratiques d'évaluation existent au niveau international (dispositifs PISA, PASEC, analyses sectorielles conduites par des organisations internationales, etc.), il n'existe que très peu de dispositifs nationaux spécifiquement organisés en ce domaine dans les pays de la CONFEMEN. Par ailleurs, la prise en compte des évaluations intégrées au fonctionnement ordinaire des systèmes éducatifs, permettant l'amélioration de la gestion des établissements sur le plan local, est relativement insuffisante. Les fortes disparités constatées sur ce plan invitent à souligner que des progrès sont à rechercher sur cet aspect du fonctionnement des systèmes éducatifs.

Encadré 3 : L'évaluation du système éducatif en France

En **France**, depuis la fin des années 80, une politique d'évaluation a été définie au niveau national. Elle vise à insuffler au sein du système éducatif tout entier une nouvelle culture de l'évaluation fondée sur la recherche de la performance. Ces évaluations se déclinent comme suit :

1- Évaluation des enseignements

L'évaluation des enseignements est du ressort de la Direction de l'évaluation, de la prospective et de la performance du ministère de l'éducation nationale (DEPP). La DEPP réalise des études sur les pratiques pédagogiques, seule ou en collaboration avec les inspections générales et avec des équipes de recherche ou des organismes extérieurs, nationaux ou internationaux.

Les inspecteurs généraux de l'éducation nationale participent à l'évaluation d'ensemble du système éducatif, à l'exception de l'enseignement supérieur. Leur évaluation porte sur les contenus d'enseignement, les programmes, les méthodes pédagogiques et les moyens mis en œuvre. Des centres de recherche sont spécialistes des questions de pédagogie, en particulier l'Institut national de recherche pédagogique (INRP).

Concernant l'enseignement supérieur, l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES) effectue des évaluations fondées sur une appréciation systématique et objective des établissements d'enseignement supérieur et de recherche. L'Agence conduit, en outre, des évaluations de portée générale, thématiques et transversales qui portent sur des sites ou des structures fédératives.

2- Évaluation des politiques éducatives

Les politiques publiques françaises, mises en place pour améliorer la réussite des élèves et les résultats du système éducatif, sont évaluées par la Direction de l'évaluation, de la prospective et de la performance (DEPP) ainsi que par des laboratoires de recherche et des instituts.

3- Évaluation de la gestion

L'Inspection générale de l'administration de l'Éducation nationale et de la recherche effectue des études sur les aspects administratifs, financiers et organisationnels, à tous les niveaux du système éducatif français. Les rapports de la Cour des comptes complètent cette évaluation de l'utilisation des fonds publics par le ministère chargé de l'Éducation nationale.

4- Évaluations diagnostiques

Elles sont des outils professionnels pour les enseignants afin de permettre l'établissement d'un diagnostic individuel des élèves à partir duquel l'enseignant définit et met en œuvre des actions pédagogiques adaptées.

5- Évaluations-bilans nationales

Un cycle d'évaluations-bilans nationales, dans la plupart des disciplines, a été mis en place en 2003 par la DEPP. Chaque année, des échantillons d'élèves passent des épreuves standardisées qui permettent d'apprécier les niveaux de connaissances et de compétences des promotions de fin d'école et de fin de collège, en référence aux programmes. Ces évaluations portent sur six domaines avec une évaluation par année d'un de ces domaines : lecture, langues vivantes, attitudes à l'égard de la vie en société, histoire, géographie et instruction civique, sciences et mathématiques. De plus, des évaluations sont mises en place en fin de cycle primaire et de collège pour répondre aux indicateurs demandés dans le cadre de la loi organique relative aux lois de finances.

6- Comparaisons internationales

Il existe de nombreuses évaluations à caractère international qui permettent de comparer la performance des systèmes éducatifs à travers le monde. Différents organismes internationaux mènent de belles études : l'IEA, l'OCDE, l'UNESCO, etc. C'est la DEPP qui participe, pour la France, aux travaux de ces organismes. Par ailleurs, il existe un réseau d'information sur l'éducation en Europe (Eurydice) dont l'unité française fait partie de la DEPP.

7- Statistiques

La Direction de l'évaluation, de la prospective et de la performance (DEPP) est chargée de la conception et de la gestion du système d'information statistique du ministère. La DEPP publie régulièrement les résultats de ses études.

Un des buts de l'évaluation est de rendre compte de l'état du système éducatif. Dans ce cadre, la DEPP assure une diffusion adaptée des résultats des évaluations auprès des élus, des journalistes, dans les milieux économiques et auprès du grand public. Un autre but est d'aider au progrès des élèves. Cela nécessite en particulier une articulation forte avec les objectifs généraux du système éducatif ainsi que les résultats. C'est pourquoi la DEPP développe des travaux visant à évaluer le degré d'atteinte des objectifs généraux du service public de l'éducation, et notamment l'évolution du capital de compétences et de connaissances détenu, d'une part, par les élèves à certaines étapes importantes du cursus et, d'autre part, par les titulaires des différents diplômes.

Encadré 4 : L'évaluation du système éducatif au Niger

Au Niger, c'est seulement en 2003, à la faveur de la mise en œuvre du Programme décennal de développement de l'Éducation (PDDE), que fut mis en place un dispositif national d'évaluation, avec pour objet d'évaluer et de suivre les acquis scolaires. Cette initiative, qui concerne l'éducation de base (primaire, 1^{er} cycle secondaire et alphabétisation), s'inscrit dans l'optique de la scolarisation primaire universelle qui, pour son suivi, nécessite des indicateurs pour l'évaluation de la qualité et du contenu de l'enseignement. Depuis 2004, des évaluations annuelles sont conduites. Elles se résument ainsi qu'il suit :

1- Évaluation des enseignements

L'évaluation des enseignements s'opère à deux niveaux : **i)** l'évaluation des prestations pédagogiques des enseignants, à travers des inspections et visites de classe, effectuées par des inspecteurs et conseillers pédagogiques en place au niveau des Inspections de l'enseignement de base (IEB), au nombre d'une cinquantaine ; **ii)** l'évaluation et le suivi des acquis scolaires que conduit, au moins une fois tous les deux ans, la Division de l'évaluation et du suivi des acquis scolaires (DESAS) de la Direction de l'évaluation, des examens et concours (DEEC) du Ministère de l'Éducation nationale. La première forme d'évaluation des enseignements porte sur les pratiques pédagogiques, la seconde, sur l'évaluation des apprentissages des élèves ou apprenants.

2- Évaluation des politiques éducatives

Dans le cadre du suivi de la mise en œuvre du PDDE, il est effectué chaque année une revue conjointe «gouvernement-PTF» du Programme décennal. Ces revues s'orientent de plus en plus vers des évaluations thématiques. À l'heure actuelle, c'est la seule forme systématisée d'évaluation des politiques éducatives au Niger.

3- Évaluation de la gestion

Au plan de l'évaluation, la dimension «gestion» ne bénéficie pas de la même attention que les autres. Pour les aspects administratifs et organisationnels, une seule évaluation a été réalisée, en 2002-2003, dans le cadre de la préparation du PDDE, afin d'asseoir le dispositif institutionnel du Programme. Pour les aspects financiers, depuis 2006, la gestion du PDDE fait l'objet d'audits externes chaque année et de contrôle parfois de la part de l'Inspection d'État.

4- Évaluations diagnostiques

Le recensement scolaire effectué chaque année donne lieu à l'établissement d'un bilan diagnostic du système éducatif, bilan réalisé par la Direction des études et de la programmation (DEP). Ce travail sert de base à la programmation et à la formulation des plans d'actions. En revanche, les évaluations diagnostiques du type outils conçus pour orienter les enseignants dans la définition et la conduite d'actions pédagogiques de remédiation, n'ont pas encore cours. Au niveau du dispositif national d'évaluation et de suivi des acquis scolaires, les mécanismes de rétroaction devant conduire à l'action ne fonctionnent pas pour l'instant.

5- Évaluations-bilans nationales

Depuis 2002, la DESAS, seule ou dans le cadre de programmes internationaux d'évaluation (PASEC et MLA notamment) réalisent, au moins une fois tous les deux ans des évaluations- bilans nationales, pour les disciplines jugées fondamentales, à savoir la lecture, les mathématiques et les sciences, pour les niveaux CP, CE2 et CM2. Depuis peu, cet exercice est étendu à la 8^e année de scolarité (la classe de 5^e des collèges). Ces évaluations- bilans nationales sont réalisées à partir de tests standardisés sur un échantillon d'élèves. Elles permettent, d'une part, d'évaluer le niveau de connaissances et de compétences des élèves de ces niveaux et, d'autre part, d'effectuer des comparaisons avec les résultats d'autres systèmes éducatifs.

6- Comparaisons internationales

Le Niger, dans le cadre de programmes internationaux d'évaluation, bénéficie, à travers la DESAS, d'appuis d'organismes tels que la CONFEMEN (pour le programme PASEC) et l'UNESCO/UNICEF (pour le programme MLA) pour conduire des évaluations à dimension sous-régionale, voire continentale. Ainsi les résultats des évaluations PASEC et MLA permettent de comparer les résultats observés au Niger avec d'autres pays africains et au-delà.

7- Statistiques

La production et la gestion d'informations statistiques relève de la responsabilité de la Direction des études et de la programmation (DEP). Celle-ci publie chaque année un annuaire des statistiques scolaires, un tableau de bord et le diagnostic de l'année scolaire considérée.

8- Première application possible du principe de l'évaluation formative

Une première application possible de l'évaluation formative aurait pour cible la lutte contre l'illettrisme. Parmi les objectifs généraux de cette initiative, se trouvent le diagnostic des compétences des élèves relatives à la lecture en début d'année scolaire et la proposition de remédiations en fonction de ces compétences évaluées.

En résumé, la pratique de l'évaluation, notamment dans son aspect de «rendre compte des résultats enregistrés», s'installe progressivement au Niger, à la faveur de la mise œuvre du PDDE. Il reste cependant à élargir son champ d'investigation à toutes les dimensions majeures du système éducatif (les programmes, les enseignements, les personnels, les dispositifs, les apprentissages, etc.) et à rendre effectif le deuxième aspect de l'évaluation, à savoir, « utiliser les résultats de l'évaluation comme outils de base au service de l'action ».

Pour ce qui est des pays du Sud, le diagnostic fait est celui d'une capacité d'autoévaluation des systèmes très limitée. Jusqu'à présent, les actions principales en matière d'évaluation ont été initiées principalement de l'extérieur.

Il ressort d'une enquête menée par la CONFEMEN, en 2005, auprès de 17 pays membres du Sud que, sur les 13 pays ayant des structures PASEC (équipes constituées lors des évaluations menées par la CONFEMEN), 11 pays disposent d'une structure nationale d'évaluation dont quatre intègrent complètement l'équipe PASEC. Dans les sept autres pays, seulement quelques membres de l'équipe PASEC sont intégrés à l'équipe nationale. Ces résultats montrent que non seulement tous les pays de la Francophonie ne disposent pas de structure d'évaluation, mais que la plupart n'ont pas de politique claire en matière d'évaluation.

L'analyse des résultats de cette enquête montre également que seuls sept pays ont réalisé des activités d'évaluation nationale de type évaluation des acquis scolaires.

2.1.2. Les évaluations nationales

Dans le cadre de l'EPT, la Banque mondiale a développé une méthodologie d'analyse et d'évaluation des systèmes éducatifs — diagnostic sectoriel — qui vise à explorer le fonctionnement du système éducatif pour mieux comprendre quelles sont ses forces et ses faiblesses; elle vise par la même occasion à identifier les sources d'inefficience et à suggérer les voies possibles de les réduire. Elle examine sous des angles variés, analytiques et factuels, un certain nombre d'aspects importants concernant le fonctionnement et le financement du système. Le Rapport d'état d'un système éducatif national (RESEN) est donc un élément d'évaluation des progrès réalisés et des difficultés rencontrées par les systèmes éducatifs en général sur un horizon de dix ans.

Encadré 5 : La méthodologie RESEN

Le **RESEN (Rapport d'état d'un système éducatif national)** désigne un type de diagnostic d'un système éducatif, produit à partir d'une démarche d'analyse précise. Cette démarche a été développée et théorisée par l'équipe Développement humain – Pays pauvres et très endettés (DH-PPTE), de la Banque mondiale, travaillant sur la région Afrique. Elle correspond à une étude en étapes, auxquelles correspondent les différents chapitres du document RESEN final :

- ▶ *Chapitre 1* : contexte global de l'évolution du secteur de l'éducation
- ▶ *Chapitre 2* : analyse des scolarisations
- ▶ *Chapitre 3* : analyse des aspects financiers
- ▶ *Chapitre 4* : efficacité interne et qualité de l'éducation
- ▶ *Chapitre 5* : efficacité externe associée à la production de capital humain
- ▶ *Chapitre 6* : équité et disparités dans les systèmes d'enseignement
- ▶ *Chapitre 7* : questions de gestion du système éducatif
- ▶ *Chapitre 8* : synthèse et estimation d'un modèle de simulation pour la politique éducative

Ces travaux ont été développés par la suite par le Pôle de Dakar et la Banque mondiale dans de nombreux pays du Sud, incluant la plupart des pays africains de la CONFEMEN.

Les travaux sur les RESEN ciblent de façon pertinente les résultats obtenus par le système éducatif en mobilisant des données d'enquêtes récentes dont :

- i) les données des structures nationales d'évaluation ainsi que celles du PASEC sur les acquis des élèves ;
- ii) les enquêtes menées auprès des ménages⁵ ;
- iii) les statistiques disponibles sur le système d'éducation.

Les RESEN ont débouché dans la plupart des pays sur la préparation de programmes décennaux de développement des secteurs éducatifs qui visent, entre autres, l'atteinte de la scolarisation universelle ainsi que l'amélioration de la gestion et du pilotage du système.

Cet exercice a permis de valoriser sur le plan national les résultats d'évaluation, notamment ceux du PASEC, en permettant leur prise en considération dans la définition des politiques de développement des secteurs éducatifs des pays concernés.

Pour les pays qui ont déjà élaboré une mise à jour du RESEN, cela leur a permis d'évaluer les progrès enregistrés dans la mise en œuvre de leur stratégie décennale et d'identifier les mesures qu'il convient de mettre en œuvre.

⁵ Les enquêtes menées auprès des ménages fournissent des informations sur la distinction des facteurs d'offre et de demande dans l'explication des scolarisations, des données sur la pauvreté qui permettent la prise en compte de la dimension pauvreté dans l'analyse des disparités scolaires et, surtout, offrent des informations importantes tant sur l'impact de l'éducation dans la sphère productive que sur une évaluation empirique du niveau et de la structure de rémunération des enseignants.

La diffusion des résultats du RESEN dans les pays a permis aux différents acteurs d'être informés sur la situation du secteur, ce qui n'a pas été toujours du goût des responsables des secteurs éducatifs (propositions suscitant parfois des réactions assez vives des responsables des secteurs éducatifs, qui se sentaient mis en cause)⁶.

De rares pays du Sud ont mis en œuvre des évaluations nationales, appuyées pour la plupart par une expertise technique extérieure. À notre connaissance, aucun de ces pays n'a réellement mis en place une véritable stratégie de diffusion des résultats de son évaluation. Par conséquent, cela fait que la dimension compte-rendu de l'évaluation n'a pas véritablement été prise en compte. Cependant, certains pays ont tenu compte des résultats dans la définition de leurs politiques éducatives.

Dans le cadre de la mise en œuvre de l'EPT, certains pays ont mis en place quelques dispositifs ou activités d'évaluation.

Encadré 6 : Exemple de Madagascar

À **Madagascar**, plusieurs actions intégrant un volet «évaluation des acquis scolaires» ont été réalisées. Le programme d'Amélioration de la gestion éducative pour l'Afrique (AGEPA) inclut une évaluation d'impact d'envergure, qui administre les tests PASEC auprès des élèves. Certains bailleurs (tels l'UNICEF et l'AFD) ont aussi intégré un volet évaluation des acquis dans leur programme (approche par les compétences, écoles pilotes, etc.). Les enquêtes sont de bien moindre envergure que celles menées dans le cadre de l'AGEPA. Dans tous les cas, il s'agit d'opérations ponctuelles qui ont pour objectifs d'évaluer l'impact des projets et initiatives des partenaires. Il ne s'agit pas d'un cycle d'évaluation récurrent et il n'existe pas de structure spécifique qui réalise ces activités à plein temps. Ce sont les techniciens de la direction de la planification, appelés à intervenir sur de nombreux chantiers, qui réalisent ces évaluations, avec l'appui technique et financier de partenaires au développement.

Encadré 7: Exemple du Sénégal

Au **Sénégal**, l'INEADE (Institut national d'étude et d'action pour le développement de l'Éducation) dispose d'une cellule d'évaluation qui mène, depuis 2000 des enquêtes appelées Système national d'évaluation du rendement scolaire (SNERS). Les résultats de ces enquêtes sont intégrés aux documents sectoriels, dans le volet qualité de l'éducation. C'est un rare cas d'enquêtes mises en place spécifiquement pour suivre l'évolution de la qualité. Cependant, les items utilisés ne sont pas toujours les mêmes entre les vagues d'enquête, et la dernière vague (2006) a été menée dans des temps records, suite à une demande de la Banque mondiale.

Encadré 8 : Exemple du Burkina Faso

Au **Burkina Faso**, l'Agence canadienne de développement international (ACDI) apporte un soutien financier à une cellule d'évaluation. L'appui est programmé en deux phases. Sous l'encadrement d'un expert international, la première phase a pour objectif le renforcement de la capacité technique de la cellule d'évaluation en matière de méthodologie et de mise en œuvre. Dans ce cadre, la cellule réalise chaque année une enquête nationale. À la deuxième phase, l'équipe met en œuvre, de façon indépendante, les activités d'évaluation, lesquelles s'inscrivent dans le pilotage du système éducatif du pays.

⁶ Les exercices RESEN ont porté au grand jour la question toujours présente mais occultée de la gestion du système. Auparavant, les analyses des systèmes éducatifs se préoccupaient plus des questions d'accès et des aspects pédagogiques.

Encadré 9: Exemple de la Mauritanie

En **Mauritanie**, les activités d'évaluation ont été confiées, en 1999, à un service d'évaluation relevant de l'Institut pédagogique national (IPN). La formation des évaluateurs a démarré à cette date, huit évaluateurs ont été formés à l'Institut de recherches en économie de l'Éducation (IREDU) de l'Université de Bourgogne (France) et à l'Université de Louvain (Belgique). En 2003, une Direction de l'évaluation a été créée sans que le service de l'évaluation de l'IPN ne soit dissout. Jusqu'à la fin 2005, ces deux structures ont fonctionné en parallèle sans concertation entre elles. Le nouvel organigramme du ministère des Enseignements fondamental et secondaire adopté en 2006 prévoit la création d'une Cellule nationale d'évaluation rattachée au Cabinet du ministre.

Depuis sa création, la structure d'évaluation a mené de nombreuses évaluations sur les programmes au fondamental, sur le contexte des niveaux 2^e, 4^e, 5^e et 6^e années du primaire et sur les acquisitions au niveau des 5^e C et D du secondaire⁷ ainsi qu'une évaluation des résultats des examens. De même, elle a piloté et mis en œuvre les enquêtes MLA et PASEC.

Il convient de souligner que toutes les évaluations menées en dehors des deux dernières, ci-haut mentionnées, l'ont été sur financement de la Banque mondiale et, qu'à cette date, aucune inscription budgétaire n'est prévue pour les activités d'évaluation. Il importe de signaler que la prise en compte de la cellule dans le nouvel organigramme est de nature à lever cette contrainte.

Par ailleurs, il faut noter que seuls les résultats des évaluations MLA et PASEC ont été diffusés. De plus, les principaux résultats de ces évaluations sont pris en compte dans le cadre de la mise en œuvre du programme décennal de développement du secteur de l'éducation.

Il apparaît que les États du Sud sont encore en retrait quant à la place de l'évaluation et de l'utilisation de ces résultats pour le pilotage des systèmes éducatifs. Il convient impérativement de renverser cette tendance pour pouvoir atteindre les objectifs de l'EPT dans un horizon raisonnable.

2.1.3. État des lieux de l'évaluation dans l'espace francophone

2.1.3.1. Une intégration relative dans l'espace francophone

Quelles que soient leur organisation administrative et les dénominations retenues, les systèmes éducatifs recouvrent un certain nombre de fonctions dont les principales sont la formation, les programmes d'études, la planification, la programmation, la gestion des ressources humaines, les statistiques, la carte scolaire, les constructions et l'équipement ainsi que les finances. En d'autres termes, il s'agit du processus d'élaboration et de mise en œuvre de politiques éducatives. Dans certains cas, ces fonctions sont confiées à des services ou à des directions.

Le questionnaire administré par la CONFEMEN montre, pour de nombreux pays membres interrogés, qu'il n'existe pas de structure spécifique consacrée à l'évaluation, ou alors il s'agit d'intitulés comme «Direction des études, de la programmation et de l'évaluation». Dans d'autres pays (Luxembourg ou France, par exemple), en revanche, la fonction d'évaluation est remplie par une direction à part entière, parfois même appuyée par des organismes de type «Agence d'évaluation de la recherche et de l'enseignement supérieur». Ces choix sont significatifs des priorités retenues et de la place effective, plus ou moins centrale ou

⁷ Sections retenues en raison d'une priorité accordée, dans le secondaire, aux enseignements scientifiques.

marginale, qu'occupe la fonction d'évaluation dans le système éducatif.

Or, ce point est essentiel, du fait des engagements pris et des contraintes de calendrier : les pays engagés sont à la fois partie prenante d'évaluations internationales (ce qui engage leur crédibilité) et aussi évaluateurs de leur propre système, notamment en ce qui concerne la fiabilité des données fournies aux décideurs ainsi que de l'efficacité de la mise en œuvre de leurs politiques éducatives.

Confier les missions d'évaluation à un service au sein d'une direction peut entraîner bon nombre d'effets indésirables. En effet, le lien hiérarchique entre le service et la direction ne garantit pas toujours l'indépendance totale des équipes chargées de l'évaluation. De plus, si une mission d'évaluation amène à s'intéresser à d'autres services ou à d'autres directions, il n'est pas du tout certain qu'un réflexe « protectionniste » ne viendra pas entraver la collecte et le traitement des données.

Une deuxième série de faits amène à se poser des questions sur le positionnement de l'instance d'évaluation dans un système éducatif. En effet, il est indispensable de s'assurer que les données collectées par les différents éléments de l'administration centrale soient fiables et de qualité. Par exemple :

- L'élaboration d'instruments statistiques a fait d'énormes progrès dans bon nombre de pays d'Afrique subsaharienne. On ne compte plus les projets ou les programmes qui comportent un volet d'appui à l'administration centrale et, plus particulièrement, à la statistique. Par exemple, l'introduction de logiciels ou de systèmes d'information pour la gestion de l'éducation (SIGE), comme ceux que met en place l'ADEA, accompagnés d'actions de formation (SISED) ont permis de réaliser de nombreux progrès. Néanmoins, il reste beaucoup à faire, comme le soulignent, parfois durement, de nombreuses publications. Par exemple, on lit dans le Cadre d'action de Dakar, « Lorsque les gouvernements ont vraiment à cœur les résultats de l'éducation, ils sont conscients de l'importance des statistiques et de la nécessité de disposer d'institutions crédibles et indépendantes chargées de les établir. Le bilan de l'EPT à l'an 2000 a révélé l'existence d'importantes lacunes dans les données ». On lit également dans un rapport de la Banque mondiale (2001) :

« Bien que des indicateurs standards sur les apports (inscriptions, professeurs et écoles) et les flux d'étudiants soient généralement disponibles, ils sont souvent peu fiables et insuffisamment détaillés. C'est ainsi que [...] sur des indicateurs essentiels pour l'évaluation 2000 de l'EPT, seul un tiers des pays d'Afrique subsaharienne était en mesure de fournir des données relatives aux inscriptions pré primaires, la moitié ne disposait pas de données spécifiques par tranche d'âge pour calculer les ratios nets d'inscription, et un quart ne disposait d'aucune donnée relative à la participation des filles ».

- L'élaboration de la carte scolaire a elle aussi bénéficié de soutiens importants et a connu des progrès considérables en Afrique subsaharienne (Sénégal, Mauritanie, Cameroun, par exemple). Ici encore, cependant, il reste beaucoup à faire. C'est dans cet esprit que la méthodologie pour la réalisation d'un Rapport sur l'état des systèmes éducatifs nationaux (RESEN) présente les méthodes d'analyse des scolarisations et, notamment, des relations entre problèmes d'accès et questions de couverture scolaire. Les auteurs analysent en particulier les questions d'offre et de demande

ainsi que leur impact sur l'accès et la fréquentation effective de l'école. Pour ce faire, ils comparent les situations dans lesquelles l'offre d'éducation est complète et celle où elle est «lacunaire» (absence d'école ou écoles n'offrant pas une scolarité complète). Ils tirent de ces analyses des enseignements sur les taux d'inscription, d'abandon ainsi que de redoublements.

Une telle approche souligne la nécessité d'intégrer dans la politique de développement de l'accès à l'école des facteurs extrêmement divers, qui touchent le plus souvent aux conditions dans lesquelles se forme et se développe (ou pas) la demande scolaire (facteurs géographiques, activités économiques locales, image de l'école, etc.). Au total, il semble avant tout «utile de développer une structure institutionnelle dans laquelle on articule (dans les modes de fonctionnement et éventuellement de financement, ainsi que dans les modes de relation entre l'État et les diverses communautés) les nécessités et les exigences d'une école nationale, d'une part, et avec les aménagements nécessaires pour rencontrer, au niveau local, les demandes effectives des familles, d'autre part»⁸.

En ce qui a trait à la diffusion, dans l'espace francophone, malgré le fait que des enquêtes nationales ou internationales ont été réalisées, on le verra plus précisément dans les études de cas, les résultats ont été peu diffusés au niveau national ou parfois contestés, parce qu'ils allaient souvent à l'encontre d'idées reçues, de conceptions éducatives traditionnelles et de pratiques administratives établies. Les équipes nationales des programmes internationaux d'évaluation peinent à présenter et à faire accepter les résultats à leurs autorités et à la société civile, en particulier les syndicats, qui ne sont pas toujours associés à la restitution. Quand la restitution des résultats a été effectuée dans de bonnes conditions, certaines réformes ont pu être conduites sur la base des résultats, notamment sur le redoublement (Bernard et Michaelowa, 2006). Si le PASEC a mis en place une véritable stratégie d'information, celle-ci est parfois absente au niveau national.

Les enquêtes internationales ont souvent des dispositifs propres de diffusion des informations et des données, qui sont régis par des conventions signées par les pays. Les conventions PASEC stipulent ainsi que le ministère ne peut pas s'opposer à la diffusion des résultats d'une évaluation, une fois le rapport validé par le Comité scientifique du programme. Les autres programmes contiennent des dispositions similaires. Ces dispositifs particuliers ne sont pas à l'origine du peu de diffusion des résultats. Souvent la non validation des enquêtes, faute du respect des standards méthodologiques, en est à la base.

2.1.3.2. Une faiblesse du pilotage par les résultats

Les résultats de l'enquête menée par la CONFEMEN, à laquelle 23 États et gouvernements membres ont répondu, sont éloquentes à propos de la relation entre la démarche d'évaluation et le pilotage. À la question concernant l'utilisation des résultats d'évaluation, les réponses sont assez diverses. Dans certains cas, ces résultats ont servi avant tout à étoffer et à nourrir un diagnostic global du système (RESEN par exemple pour le Togo et le Niger) ou encore soutenir le dialogue avec les partenaires (tel est le cas de Maurice).

Lorsque des décisions sont citées au titre des conséquences de l'évaluation, il est fait mention de décisions «politiques» de grande ampleur : réformes des curricula, orientation de la planification stratégique et réforme de la politique des langues (12 réponses en ce sens).

⁸ Citation tirée du document de travail présenté à la 51^e session ministérielle de Maurice.

Dans deux cas seulement, les décisions décrites appartiennent à la sphère des pratiques et du pilotage : les Seychelles et le Canada/Nouveau-Brunswick, ce dernier gouvernement étant apparemment le seul cas, parmi les répondants, pour lequel les établissements d'enseignement sont instruits à conduire des actions de remédiation suite aux résultats des évaluations, sous la supervision des conseils d'éducation de district.

2.1.3.3. Des dispositifs nationaux d'évaluation insuffisamment structurés

À l'observation des résultats de l'enquête précitée, les dispositifs nationaux d'évaluation apparaissent avant tout comme extrêmement diversifiés. Ce point est notable, en ceci que ce n'est pas le cas pour les dispositifs qui répondent à d'autres fonctions, plus classiques, de l'administration scolaire. Ainsi, le suivi des activités statistiques est très généralement le fait d'une direction spécifique du ministère (20 pays sur 23) : pour remplir cette fonction, il y a une convergence forte entre les États et gouvernements quant au mode le plus adéquat de structuration administrative. Pour l'évaluation, les structures relèvent soit des corps d'inspection, soit des services ou directions des examens ou de la planification, soit de cellules *ad hoc* situées en marge de la hiérarchie, sans qu'on puisse déterminer un mode classique ou dominant d'organisation.

De plus, la moitié des États et gouvernements (11) citent plusieurs structures chargées de l'évaluation au plan national (inspection, service des examens, planification, cellule *ad hoc*, agence de la qualité, etc.), leur nombre pouvant aller jusqu'à cinq, à tel point d'ailleurs que dans certains cas, il devient probablement difficile de parler de «dispositif». D'ailleurs, le besoin de structurer la fonction d'évaluation, de l'associer à un organe stable, est fréquemment cité, notamment par Madagascar, le Congo, Maurice, les Seychelles et le Gabon, dans les suggestions faites pour concourir à l'amélioration des dispositifs.

Au-delà des dispositifs, on constate une instabilité du personnel en charge des opérations d'enquête, même quand il s'agit d'évaluations menées par des programmes internationaux. On peut légitimement se demander pourquoi, dans un contexte de rareté des personnels dotés de compétences nécessaires, on ne les a pas rassemblés au sein d'une cellule ou d'une structure pérenne en charge des activités d'évaluation, qu'elles soient nationales ou internationales.

De plus, ces personnels réalisent généralement tout ce qui a trait aux opérations statistiques, à la planification et à la carte scolaire, aux réalisations des différents rapports sectoriels avec les bailleurs, etc. Cela prouve qu'il n'y a pas de personnel qui se consacre exclusivement aux questions d'évaluation. Ces dernières sont donc prises en charge par ces personnels autant qu'ils peuvent. L'insuffisance quantitative et qualitative (notamment en matière d'évaluation) de ces personnels fait que les opérations d'évaluations restent ponctuelles.

2.1.3.4. Des données incomplètes ou non disponibles

Le questionnaire adressé aux États et gouvernements ne permet pas de cerner la réalité de la qualité et de la disponibilité des données utiles pour l'évaluation. En tout cas, aucune suggestion n'est avancée dans ce domaine en réponse à la dernière question qui concerne cet aspect.

Pour autant, il apparaît, d'une part, que la participation à des enquêtes internationales, pour importante qu'elle soit, ne produit pas suffisamment de séries de données récurrentes, mais intervient seulement de façon intermittente et assez espacée ; et, d'autre part, aucun relais dans la récurrence n'est assuré par les dispositifs nationaux. Toutefois, plusieurs pays produisent des annuaires statistiques qui renseignent, notamment, sur les écoles, les effectifs et leur répartition, l'environnement scolaire, les enseignants et les résultats des examens et des concours. Ces informations permettent ainsi au ministère de disposer d'un tableau de bord de la situation statique de leur système éducatif.

2.1.3.5. Une faible implication des acteurs

Il faut noter que si les structures ministérielles et les experts voient d'un très bon œil la participation de leur pays à des dispositifs internationaux, ce n'est pas toujours le cas des syndicats d'enseignants, aussi bien au Sud qu'au Nord. Ainsi, le programme PISA ne va intégrer qu'à partir de 2009 un questionnaire destiné aux enseignants, dans son dispositif de recueil d'information. Notons que des syndicats menacent parfois de grèves si leur pays participe au programme.

Les revues sectorielles, formes de co-évaluation, organisées annuellement, constituent un bilan où chaque année, les cadres du ministère discutent des progrès accomplis (en fonction des indicateurs) et des difficultés du secteur avec les partenaires sociaux (syndicats, associations de parents d'élèves) et les partenaires techniques et financiers (bailleurs, ONG). Ces revues, qui sont un prolongement de l'initiative de Mise en œuvre accélérée de l'Éducation pour tous (IMOA-EPT), restent dans le champ international. En effet, elles président souvent au choix du volume de financement octroyé à tel ou tel pays par un bailleur.

On a vu que la fonction de compte-rendu était une dimension essentielle de l'évaluation. Or, les associations de parents d'élèves ne sont pas toujours associées aux séminaires de restitution et ne connaissent généralement pas les résultats des évaluations, autrement que par la presse. Pourtant, les communautés et les parents restent parmi les principaux bailleurs de l'éducation ainsi que les premiers concernés. Des initiatives novatrices existent en la matière, tel le programme AGEPA (Amélioration de la gestion de l'éducation dans les pays Africains) qui entend produire des outils locaux de pilotage et insister sur la diffusion d'informations auprès des communautés. Ce projet reste, pour l'instant, une initiative internationale, en l'occurrence, de la Banque mondiale.

La faible implication des syndicats et autres acteurs de la société civile, dans ces processus, induit parfois des réticences de leur part, fait qui a été souligné plus haut.

Ce faible degré d'implication est à mettre en relation avec les activités qui forment l'essentiel de la tâche dévolue aux organes chargés de l'évaluation. Les réponses au questionnaire ne laissent aucun doute à ce sujet : seules des activités de recueil de données (sur des performances d'élèves, généralement) et d'analyse sont citées par les États et gouvernements (22 cas sur 23 à l'enquête, dont un seul cite le développement de plans locaux de remédiation). La diffusion des résultats et diagnostics existe certainement mais n'est pas centrale dans la fonction d'évaluation telle que les pays la décrivent, et encore moins en ce qui concerne le suivi ou l'accompagnement des acteurs dans leur appropriation ou la définition de rétroactions.

L'accent étant ainsi mis d'abord sur la collecte et l'analyse, et très secondairement sur l'action ou la mise en perspectives des résultats, il se peut que la fonction d'évaluation apparaisse aux partenaires de l'école comme une affaire de spécialistes ou une production documentaire et administrative. Par conséquent, cela n'appelle pas à une implication forte de l'ensemble des acteurs, sauf au moment de contester des constats trop saillants ou trop contre-intuitifs.

2.1.3.6. Des lacunes dans la formation des acteurs et contraintes

Malgré les efforts de formation des évaluateurs nationaux réalisés par la CONFEMEN dans le cadre du PASEC, les capacités d'évaluation des pays restent très en deçà de ce qui est nécessaire pour bâtir un véritable dispositif d'évaluation.

De même, les ressources nationales allouées à l'évaluation ne sont pas souvent sécurisées, c'est-à-dire inscrites dans un chapitre défini du budget de l'État, et ne sont pas suffisantes, quand elles existent, pour permettre la réalisation d'évaluations fiables et valides.

Comme déjà souligné, la plupart des évaluations nationales réalisées sont, le plus souvent, non seulement commanditées par des partenaires extérieurs mais aussi financées par eux. Ces évaluations nationales sont confrontées le plus souvent à des obstacles qui empêchent leur prise en compte dans le processus de décision. Il s'agit :

- de problèmes de fiabilité et de disponibilité des données statistiques ;
- de résultats des évaluations non disponibles aux acteurs du système ;
- d'insuffisance d'implication des principaux acteurs dans le processus d'évaluation ;
- d'absence d'une culture d'évaluation à tous les niveaux, dans plusieurs pays;
- d'insuffisance de mécanismes d'autoévaluation dans les structures de gestion scolaire ;
- d'absence de recherche de performance au niveau de la gestion publique (justification des besoins d'évaluation).

Il convient donc de souligner que, si les dispositifs nationaux doivent également s'interroger sur les questions précises posées par les évaluations internationales et surtout trouver des pistes d'amélioration, il est nécessaire que ceux-ci soient pensés en partant de l'échelon local — celui de l'établissement — là où se jouent les objectifs d'accès, de rétention, de qualité et d'équité.

Les réponses au questionnaire administré auprès des États et gouvernements membres de la CONFEMEN en portent le témoignage : dans 18 cas sur 23, les ressources humaines affectées à l'évaluation sont estimées insuffisantes, en nombre et en formation. D'ailleurs, l'insuffisance numérique du personnel affecté aux dispositifs d'évaluation est fortement corrélée dans les réponses à la pénurie d'agents ayant reçu une formation convenable dans le domaine, plus qu'à un manque généralisé de main-d'œuvre.

Quant aux partenaires de l'école et de la société civile, il semble bien également que le défaut de formation puisse leur rendre peu accessibles les diagnostics issus des opérations d'évaluation. Si ce problème relève pour une part des stratégies de communication (voir plus bas), on peut imaginer que des formations-sensibilisations pourraient apporter leur contribution aux processus participatifs d'évaluation.

2.2. LES PROGRAMMES INTERNATIONAUX D'ÉVALUATION

À la 51^e session ministérielle, les ministres en charge de l'éducation ont reconnu que la comparaison de leurs différents systèmes éducatifs n'a été rendue possible que grâce aux programmes internationaux d'évaluation. Il s'agit, notamment, de TIMMS, PISA, SACMEQ, MLA et PASEC.

2.2.1. Rôles et fonctions des programmes internationaux d'évaluation

Différents types de programmes sont mis en œuvre de par le monde, certains sont destinés aux pays du Nord, d'autres à ceux du Sud. Il n'existe pas à l'heure actuelle un programme ou dispositif international francophone qui couvre à la fois les pays du Sud et ceux du Nord. Ce qui nous intéresse, ce ne sont pas les programmes eux-mêmes, bien qu'ils se substituent parfois aux dispositifs nationaux, mais bel et bien les spécificités de la participation des pays francophones à ces programmes.

Plusieurs programmes majeurs retiennent notre attention. Une récente publication de l'Institut international de planification de l'éducation (IIPE, 2006)⁹ fournit un ensemble d'informations sur ces programmes, à la fois sur les volets technique et politique, ainsi que sur les aspects liés à la communication des résultats. Cette publication est issue d'un groupe de chercheurs réunis en 2005 sous l'égide de l'UNESCO.

Association internationale pour l'évaluation des acquis scolaires (IEA)

Dans les années 1960, l'éducation comparée se développe et un groupe d'experts, rassemblés sous l'égide de l'UNESCO, crée l'IEA¹⁰. Cette association de scientifiques, à but non lucratif, réalise dès les années 1960 les premières enquêtes internationales d'évaluation des acquis scolaires, dans les pays du Nord. Les bases méthodologiques des études à vaste échelle sont dès lors jetées. À ce jour, l'IEA réalise toujours ses enquêtes (TIMSS, PIRLS), qui constituent des références scientifiques incontournables et qui ont inspiré d'autres initiatives.

Le Programme international de recherche en lecture scolaire (PIRLS)¹¹ évalue les compétences des filles et des garçons de neuf ans (élèves de 4^e année) par rapport à la lecture de textes littéraires et informatifs authentiques.

Le *Trends in International Mathematics and Science Study* (TIMSS) étudie par comparaisons internationales les acquisitions scolaires en mathématiques et en sciences. Cette étude a été menée en 1995, 1999 et 2003.

L'IEA mobilise pour ses travaux son Assemblée générale, son réseau de coordonnateurs nationaux, son propre centre d'expertise (situé à Hambourg, en Allemagne) ainsi que des centres de recherche et universités, notamment l'*Australian Council for Education Research*, en Australie, le *Boston College* et le *National Council for Education Research* aux États-Unis. Des coordonnateurs nationaux sont désignés dans chaque pays pour suivre et rendre compte du déroulement des enquêtes tout en participant à leur planification et à leur

⁹ www.iipe.unesco.org

¹⁰ www.iea.nl

¹¹ Le PIRLS a été administré pour la première fois en 2001. En 2006, 43 pays ont participé à l'évaluation.

adaptation au contexte national. Les activités de l'IEA émanent avant tout de chercheurs, sans être forcément une demande des autorités politiques nationales, ce qui pose parfois problème au moment de la prise en compte des résultats dans la formulation des politiques éducatives.

Ces enquêtes nécessitent un haut niveau de technicité : les standards liés aux taux de réponse et à la qualité des données sont élevés. Certaines conditions président à l'admission au sein de l'IEA et il serait souhaitable que les pays francophones de l'Afrique subsaharienne puissent y participer progressivement. Quelques pays francophones du Nord y participent, mais leur influence sur les choix techniques et les thèmes des études reste limitée, même s'ils collaborent entre eux, notamment lorsqu'il s'agit d'adapter les instruments en français.

Programme international pour le suivi des acquis des élèves (PISA)

Si les études de l'IEA restent une référence, l'attention des scientifiques, des politiques et surtout des médias s'est détournée ces dernières années vers le programme PISA de l'OCDE. Dans les années 1990, les États membres de l'OCDE ont créé les réseaux d'experts INES (*International Networks on Education Statistics*), chargés de réaliser des comparaisons des systèmes éducatifs à partir d'indicateurs. Il ne s'agissait pas au départ d'un instrument d'évaluation et les indicateurs avaient pour but de comparer la mobilisation des ressources, les salaires des enseignants, les niveaux de scolarisation et la qualité de l'éducation (au travers de tests).

Le programme PISA, qui évalue les compétences des élèves de 15 ans sur une base tri-annuelle, est devenu, au sein de ce réseau, une référence mondiale et les résultats font l'objet de nombreux articles de presse. Ce programme a entraîné des changements majeurs dans les systèmes éducatifs en Europe, puisque les résultats sont présentés sous la forme de palmarès ou de classement et suscitent de vives réactions, voire une véritable prise de conscience dans certains pays. De nos jours, les budgets PISA sont colossaux et des centaines de chercheurs travaillent sur ces données qui font l'objet de nombreuses publications et travaux de recherche. Le dispositif est si populaire que des jeux télévisés s'en inspirent. Des pays adoptent les curricula de pays homologues qui obtiennent les meilleurs résultats, quasiment sans adaptation au contexte national. Cependant, l'engouement pour les résultats des enquêtes PISA n'est pas véritablement partagé par tous les pays de l'OCDE et ces travaux restent parfois sous-exploités dans certains pays francophones du Nord comme la France. En effet, on constate que si les indicateurs de l'OCDE et de PISA sont intégrés dans les publications nationales des pays francophones, les données ne font pas systématiquement l'objet de rapports de recherche spécifiques ou d'analyses secondaires. De plus, une fois encore, les centres de recherche mobilisés (on parle d'un consortium PISA) n'appartiennent pas à l'espace francophone.

On notera que le niveau de productions de la recherche francophone en matière d'évaluation est très faible comparé aux milliers d'articles, revues et séminaires en langue anglaise. On retiendra donc que si les pays francophones du Nord participent aux enquêtes internationales, leur participation est plus irrégulière que celle d'autres pays, la valorisation de leur expertise technique reste insuffisante et les résultats sont peu diffusés et sous-exploités à des fins de pilotage du système éducatif. Il est difficile d'évaluer dans quelle mesure la participation des pays du Nord à ces dispositifs ou programmes internationaux a été l'occasion d'un renforcement des capacités nationales, mais on peut faire l'hypothèse

qu'elle y a contribué.

Dans les pays en développement, la réflexion sur la mise en place d'outils internationaux d'évaluation s'est engagée après le sommet de Jomtien (1990) sur l'Éducation pour tous. Dans le fond, les enjeux sont les mêmes que dans les pays du Nord : stimuler la recherche et la pratique de l'évaluation ainsi qu'améliorer le pilotage et l'efficacité des systèmes éducatifs.

Trois programmes retiennent notre attention, le SACMEQ, le MLA et le PASEC.

Southern Africa Consortium for Monitoring Educational Quality (SACMEQ)

Le programme SACMEQ a été créé en 1991 à l'initiative des ministres de l'Afrique Australe. Organisé en Consortium, ce programme est piloté par un Conseil (*Board*) des ministres et bénéficie du soutien technique de l'Institut international de planification de l'Éducation. Il couvre l'Afrique anglophone australe mais apporte aussi des appuis ponctuels à certains pays d'Asie du Sud-est (Cambodge et Vietnam).

Ses objectifs sont le suivi de la qualité par l'évaluation des niveaux de compétences des élèves, pour lesquels des seuils ont été définis. Le fonctionnement du programme, l'organisation des vagues d'enquête ainsi que la méthodologie sont inspirés des enquêtes de l'IEA. La comparaison internationale ne s'est mise en place qu'à partir de 2004, les ministres étant très réticents au départ.

Le programme accorde une grande place au renforcement des capacités et privilégie la coopération Sud-Sud. Les opérations de terrain sont réalisées par des équipes nationales (constituées de chercheurs) et l'expertise du Nord, concentrée à l'IIPE, intervient seulement dans la phase d'analyse des données. Le SACMEQ a réalisé deux vagues d'enquête internationale et près d'une trentaine de rapports nationaux, qui sont des commandes des ministres. Les résultats des études ont suscité des réformes dans les systèmes éducatifs et ont été repris dans de nombreuses publications et travaux scientifiques. Il nous est difficile de caractériser, faute d'information, la relation qu'entretient le SACMEQ avec d'éventuels dispositifs nationaux d'évaluation, mais le programme a sûrement contribué à diffuser la culture de l'évaluation dans les pays anglophones d'Afrique australe.

Monitoring Learning Achievement (MLA/UNESCO)

Le programme MLA de l'UNESCO, créé en 1992 mais suspendu depuis 2004, couvrait toutes les régions du globe et ciblait les pays en développement. Il avait pour objectifs le suivi de la qualité et le renforcement des capacités en matière d'évaluation. Il évaluait les performances d'enfants de 4^e année de scolarité primaire sur la base d'un ensemble de compétences définies à l'échelon national en compétences de la vie courante, lecture et calcul. Ce programme reposait quasi exclusivement sur une coopération Sud-Sud.

Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC)

Le PASEC, créé en 1991 par la CONFEMEN, poursuit d'autres objectifs que le SACMEQ. Il s'agit, en priorité, de rechercher les facteurs d'efficacité des politiques éducatives en relation

avec la qualité. La méthodologie, l'ancrage institutionnel, le fonctionnement du PASEC, diffèrent sensiblement de ceux du SACMEQ.

La CONFEMEN a fixé au PASEC quatre objectifs :

- identifier des modèles d'écoles efficaces et peu coûteux en comparant, à l'échelle nationale et internationale, les performances des élèves, les méthodes d'enseignement et les moyens mis en œuvre ;
- développer, dans chacun des États participants, une capacité interne et permanente d'évaluation du système éducatif ;
- diffuser librement les résultats obtenus, de même que les méthodes et les instruments d'évaluation préconisés ;
- renforcer le rôle d'observatoire permanent des systèmes éducatifs du Secrétariat technique permanent de la CONFEMEN.

Le PASEC a connu des évolutions notables. Dans une première phase, les recherches étaient entièrement pilotées par des centres de recherche du Nord. Puis, le Secrétariat technique permanent (STP) de la CONFEMEN a recruté plus de personnel et les rapports ont été réalisés à son niveau, à partir de l'année 1995. L'expertise du Sud a été largement mobilisée, au niveau des équipes nationales, responsables des opérations de terrain, puis par le recrutement de personnel au STP : conseillers techniques, contractuels et stagiaires.

Outre des évaluations nationales (21 pays ont à ce jour été évalués ou sont en cours d'évaluation), le PASEC réalise des études thématiques (par exemple sur le redoublement, la formation des maîtres, etc.) et contribue au renforcement de capacités dans les pays participants. Il n'a cependant pas vocation à répondre à tous les besoins en évaluation ressentis à l'échelle d'un système éducatif national.

En somme, après 17 ans d'existence, le bilan du PASEC est significatif :

- plus de 24 rapports d'évaluations dans 21 pays d'Afrique ;
- deux publications et une autre en cours de rédaction ;
- 100 missions de formations ;
- 12 ateliers internationaux de formation ;
- reconnaissance dans la communauté scientifique internationale ;
- développement de plusieurs partenariats (IEA, UNESCO-BREDA).

Les équipes nationales sont constituées de cadres des ministères de l'éducation mais n'appartiennent pas au monde de la recherche. Toutefois, la coopération du Sud reste largement insuffisante. Bien que le souhait de la CONFEMEN soit de promouvoir cette approche¹², plusieurs raisons expliquent qu'elle ne soit pas réellement mise en œuvre. En effet, pour pouvoir réaliser ce type d'opérations, il faut que les pays disposent d'une masse critique de spécialistes en évaluation des acquis scolaires et qu'ils soient disponibles pour accomplir des missions ou des appuis.

¹² Voir les Actes de la 52^e session ministérielle de la CONFEMEN de Niamey (Niger, 2006)

Le PASEC et le SACMEQ qui concernent respectivement les pays francophones et anglophones, complémentaires dans leurs missions respectives, pourraient développer des synergies prometteuses en vue de fournir des informations plus complètes pour l'élaboration et la mise en œuvre des politiques éducatives, notamment dans les pays bilingues au sein desquels ces programmes interviennent.

Concernant l'existence ou non d'une structure d'évaluation des acquis scolaires, les réponses au questionnaire CONFEMEN révèlent que seuls huit pays sur 23 répondants disposent d'une équipe PASEC et d'une cellule d'évaluation. Les pays qui réalisent des évaluations nationales ne se reposent pas toujours sur l'équipe PASEC en place quand bien même ses membres ont participé à des ateliers de formation ou ont été formés par le STP sur place ou à Dakar. Ceci est le symptôme d'un positionnement institutionnel inadéquat des structures, quand elles existent, et d'une faible valorisation des personnels en charge de l'évaluation.

Les enquêtes internationales ne peuvent pas seules assurer la diffusion de la culture de l'évaluation, encore faut-il que du personnel et des structures pérennes existent et que des stratégies soient mises en place pour asseoir les pratiques de l'évaluation, une fois les enquêtes internationales achevées. Il semblerait donc que malgré les efforts entrepris par diverses organisations, les pays du Sud n'aient pas encore tous atteint la masse critique de spécialistes en évaluation nécessaire.

2.2.2 La coopération entre les programmes internationaux d'évaluation

La coopération (de manière formalisée) entre les programmes internationaux était, jusqu'en 2005, inexistante. L'IEA et l'OCDE entretiennent peu de relations¹³. La collaboration entre le PASEC et le SACMEQ n'était pas non plus suivie ; il en est de même avec le MLA.

En 2005, la Banque mondiale a mis sur pied un programme appelé «Partenariat global pour l'évaluation des acquis scolaires», dont la stratégie se situe à trois niveaux :

- augmenter la participation des pays en développement aux enquêtes internationales TIMSS et PIRLS ;
- renforcer les programmes régionaux ;
- appuyer les structures d'évaluation nationales.

Le PASEC, le WAEC, le SACMEQ et PREALC¹⁴ participent à ce programme, qui est mis en œuvre par l'IEA. Une première réunion de coordination a eu lieu en octobre 2006 pour s'entretenir sur les modalités d'appui de la Banque mondiale et sur d'éventuelles collaborations d'ordre technique, notamment sur le rapprochement des différents tests utilisés par les enquêtes internationales.

En dehors des financements substantiels que pourrait amener ce programme sur dix ans, des partenariats techniques peuvent être envisagés entre les différentes institutions, notamment sur ces différents éléments :

¹³ Voir le rapport de l'Assemblée générale de l'IEA 2005. www.iea.nl

- méthodes d'évaluation des besoins en matière de renforcement des capacités;
- constitution de banques d'items (ancrage des enquêtes) ;
- développement d'outils de saisie et de traitement de données ;
- séminaires internationaux sur les méthodes d'évaluation.

Les programmes régionaux pourraient servir à catalyser les différents appuis techniques en tenant compte des spécificités institutionnelles, culturelles et linguistiques. L'expertise technique de l'IEA pourrait être mise à profit.

¹⁴ Le WAEC (*West African Examination Council*) et le PREALC (Programme régional d'emploi pour l'Amérique Latine et les Caraïbes) n'ont pas pour objet principal l'évaluation standardisée et comparée des acquisitions scolaires, mais sont néanmoins intéressés à ce projet de la Banque mondiale.

CHAPITRE III :

LES ORIENTATIONS

Face à la situation actuelle en matière d'évaluation, telle que présentée dans le chapitre précédent, il est clair que le renforcement ou la mise en place d'organes et d'instruments de pilotage apparaît indispensable.

C'est d'ailleurs l'une des préoccupations majeures du Cadre d'action de Dakar (2000), qui insiste sur ce point dans sa partie IV consacrée aux stratégies¹⁵: «Assurer un suivi systématique des progrès accomplis pour atteindre les objectifs et mettre en œuvre les stratégies de l'EPT aux niveaux national, régional et international», et qui insiste fortement sur le développement élargi de l'évaluation : «surveiller les performances et évaluer les résultats qualitatifs et quantitatifs», «Fournir des données aux responsables de la planification, de la gestion et de l'évaluation», «les progrès accomplis [...] doivent être régulièrement et systématiquement évalués».

Cette insistance s'inscrit dans une tendance générale qui valorise de plus en plus l'évaluation et qui lui accorde, dans le pilotage de systèmes de plus en plus complexifiés, une place centrale et un rôle essentiel.

3.1. LE MÉMORANDUM SUR L'ÉVALUATION DES SYSTÈMES ÉDUCATIFS

Lors de la 51^e session, tenue en octobre 2004 à Maurice, les ministres de la CONFEMEN ont adopté le Mémoire sur l'évaluation des systèmes éducatifs¹⁶. Ils ont relevé que les défis et les enjeux de l'Éducation pour tous adoptés au Forum de Dakar en avril 2000, visant particulièrement la scolarisation primaire universelle en 2015 et l'élimination de la disparité entre genres, nécessitent tout à la fois la mobilisation de ressources additionnelles substantielles que l'amélioration de l'efficacité avec laquelle elles pourront être utilisées. Pour l'atteinte de ces objectifs, l'adoption de politiques éducatives plus efficaces de même que leur mise en œuvre concrète pour assurer une meilleure transformation des ressources en résultats s'impose.

Dans cette perspective, les ministres considèrent que l'institution de l'évaluation sous ses différentes formes permettra, d'une part, d'identifier les voies par lesquelles ces gains d'efficacité pourront être obtenus et, d'autre part, de rendre compte à la communauté nationale et internationale des progrès réalisés.

L'analyse montre que si des pratiques d'évaluation existent au niveau international (dispositifs PISA, PASEC, analyses sectorielles conduites par des organisations internationales), il n'existe que très peu de dispositifs nationaux spécifiquement organisés en ce domaine. Par ailleurs, la prise en compte des évaluations intégrées au fonctionnement ordinaire des systèmes éducatifs, permettant l'amélioration de la gestion des établissements sur le plan local, est relativement insuffisante. Les fortes disparités constatées sur ce plan invitent à souligner que des progrès sont à rechercher sur cet aspect du fonctionnement des systèmes éducatifs.

Sur la base de ce diagnostic, les ministres de la CONFEMEN et chefs de délégation, ont convenu des principes suivants :

- le caractère essentiel de l'évaluation pour le pilotage des systèmes éducatifs ;

¹⁵ Les citations sont situées au niveau de la stratégie 11, à la page 21 du Cadre d'Action de Dakar, 2000.

¹⁶ Le présent texte du Mémoire est tiré des Actes de la 51^e session ministérielle de la CONFEMEN.

- la nécessité d'intégrer la fonction d'évaluation au cœur de ces systèmes ;
- le besoin d'articuler les politiques nationales d'évaluation aux initiatives et dispositifs internationaux ;
- l'indispensable inscription de l'évaluation dans une démarche transparente, en vue d'une meilleure gouvernance et de l'implication de tous dans la mise en œuvre des réformes de politiques éducatives.

Compte tenu de ces principes, ils ont adopté les recommandations suivantes :

- faciliter la mise en place d'un dispositif national d'évaluation ayant un ancrage institutionnel adapté et convenablement doté en personnels et en moyens ;
- développer, articuler et valoriser ces dispositifs nationaux par des actions concertées de formation et par l'élaboration d'instruments, afin que les travaux engagés soient d'une qualité technique appropriée et qu'ils permettent une comparabilité internationale des résultats obtenus ;
- promouvoir de nouvelles modalités de gestion sur le plan local fondées sur une démarche de pilotage par les résultats dans le cadre d'une dynamique partenariale engageant l'ensemble de la communauté éducative. Un tel processus impliquera à la fois une activité d'information/ formation de ces différents acteurs de même que la mise à disposition des données pertinentes pour l'animation et la supervision d'ensemble ;
- identifier les aspects opérationnels de ces actions en examinant les options possibles, les outils nécessaires, les activités concrètes à mettre en œuvre, les calendriers et les budgets adéquats.

En reconnaissant que les dispositifs d'évaluation qui existent à l'échelon international sont bien développés et que les constats tirés interpellent fortement et directement les pays, la CONFEMEN a réitéré l'importance de compléter ce niveau d'évaluation par des dispositifs nationaux plus ciblés afin que ces dispositifs nationaux puissent répondre aux questions précises posées par les évaluations internationales et surtout trouver des pistes d'amélioration possibles.

Il convient aussi de savoir qu'on ne peut pas tout évaluer tout le temps dans un système d'éducation, c'est pourquoi il importe d'identifier les aspects prioritaires qui doivent faire l'objet d'une évaluation régulière. On sait que l'évaluation est tributaire de l'information disponible, il est donc nécessaire, une fois les objectifs d'évaluation définis, d'identifier les informations disponibles et les besoins en information. La mise en place de l'évaluation dans un système éducatif ne se limite pas à la production de quelques indicateurs. Elle implique un processus plus complexe, notamment sur le plan institutionnel, qui doit déboucher sur un véritable système d'évaluation.

3.2. PRINCIPE DIRECTEUR : DÉVELOPPEMENT ET PARTAGE D'UNE CULTURE D'ÉVALUATION

La dissémination d'une culture de l'évaluation signifie que cette démarche est appelée à irriguer tous les aspects de la gestion quotidienne du système éducatif, en devenant la première étape du pilotage administratif local, régional et national des systèmes, par la mise en œuvre très rapide de rétroactions concrètes et locales en réponse aux constats des évaluations régulières. Elle signifie également que le développement des systèmes éducatifs en direction des objectifs de l'EPT doit s'appuyer principalement sur une élaboration empirique des mesures stratégiques, en relation étroite avec les observations

régulières opérées sur les résultats obtenus.

Au delà, la CONFEMEN a un rôle spécifique à impulser et à jouer dans le champ de la Francophonie. En effet, par ses origines comme par ses développements contemporains, la pratique de l'évaluation des systèmes éducatifs doit beaucoup au monde anglophone, et reste donc très marquée par des approches de cette sphère culturelle qu'on a peut-être tort d'opposer trop radicalement à la Francophonie. Certes, les traditions scolaires sont spécifiques, cependant les finalités, elles, sont largement partagées.

La définition des performances attendues et mesurées, la mise au point des outils correspondants sont autant de chantiers certes techniques mais également marqués par les espaces culturels auxquels l'évaluation doit s'appliquer. Le monde francophone peut et doit, en investissant ces problématiques, développer et s'approprier les méthodes d'évaluation qui, tout en permettant une bonne échelle de comparaison, rendent compte aussi des traits dominants de sa culture scolaire.

3.3. LES PRINCIPES ET OBJECTIFS DES DISPOSITIFS NATIONAUX D'ÉVALUATION

Déjà à l'heure actuelle, de nombreuses données utiles sont disponibles dans les systèmes éducatifs ; il importe qu'elles soient mises en forme et mises en relation au bénéfice de l'amélioration du fonctionnement et de l'efficacité des systèmes. De plus, une meilleure définition des responsabilités des différents acteurs de l'éducation, et ce, à tous les échelons, permettra de mieux transformer ces informations en actions. L'enjeu pour plusieurs États et gouvernements francophones est donc de créer, à partir d'une situation d'information et de répartition des rôles incomplète, un véritable système d'évaluation centré sur les enjeux premiers de la scolarisation selon les États et gouvernements.

À cette fin, les principes et objectifs de tout dispositif national d'évaluation consisteront à :

- favoriser un pilotage par les résultats en produisant des informations sur les résultats obtenus, les ressources, les processus et les contextes ;
- articuler l'information aux différents échelons du système, mais en privilégiant l'échelon local car c'est au niveau des écoles que les résultats sont produits ;
- mettre en rapport les informations produites avec des grandeurs de référence (résultats d'autres écoles, moyennes de circonscription, de région, nationales ou internationales) afin de repérer les situations anormales et/ou les priorités d'amélioration ;
- cibler l'information pertinente pour chaque acteur, à chaque échelon, en fonction de ses responsabilités d'action ou de contrôle ;
- partager cette information, à chaque échelon du système (tant sur le plan vertical qu'horizontal), de manière transparente, au bénéfice de tous les partenaires de l'école ;
- utiliser des informations évaluatives à des fins décisionnelles.

Pour autant, il n'est pas besoin d'attendre la mise en place d'un système complet d'évaluation pour procéder, au plan national, à des opérations peu coûteuses, mobilisant des informations

déjà disponibles et permettant des analyses utiles pour le pilotage des systèmes. C'est le cas de l'exploitation des résultats aux examens nationaux à des fins de gestion pédagogique du système.

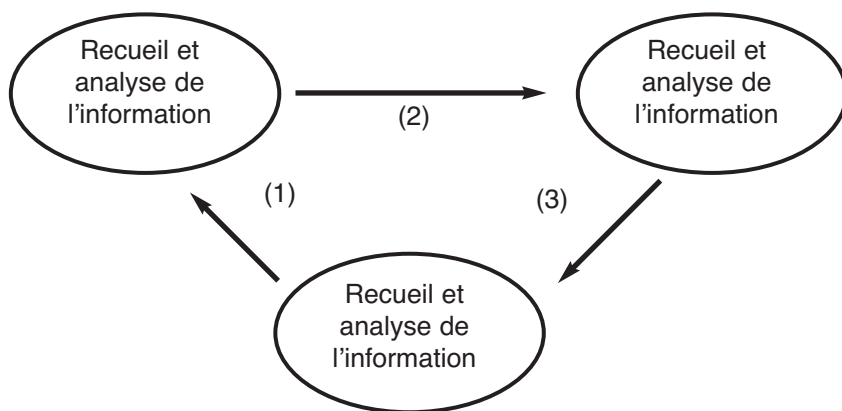
3.4. LES VOIES D'OPÉRATIONNALISATION

L'évaluation n'a de sens que lorsqu'elle devient un outil de pilotage à part entière. Il faut donc concevoir l'évaluation comme un système qui vise à l'amélioration continue du système éducatif.

3.4.1. L'institutionnalisation de l'évaluation

Le fonctionnement d'un système d'évaluation dépend étroitement du traitement de l'information comme l'illustre le schéma suivant. L'action éducative (1) produit un certain nombre d'informations dont certaines, une fois analysées, deviendront des indicateurs de fonctionnement qui pourront attirer l'attention des décideurs sur tel ou tel problème (2). Le décideur pourra alors prendre des mesures (3) selon un processus plus ou moins complexe et plus ou moins participatif pour surmonter le problème identifié.

Schéma : Circulation de l'information au sein du système d'évaluation¹⁷



Derrière ce schéma simple de circulation de l'information se cache une réalité institutionnelle complexe. Aucun des pôles représentés sur ce schéma ne représente une unité institutionnelle. Cela implique qu'à la complexité de l'information et de sa collecte, il faut ajouter celle des différents maillons qui vont intervenir dans la circulation de l'information. Il suffit de s'imaginer le chemin que doit parcourir une information sur la situation dans une classe pour atteindre un responsable du système éducatif pour toucher du doigt toute la difficulté que revêt une telle entreprise. Cela est d'autant plus difficile que l'information n'a rarement qu'un seul utilisateur et que bien souvent elle sert à des usages différents selon

¹⁷ Ce schéma a fait l'objet d'une présentation d'un rapport d'appui à Madagascar. Présentation faite par M. Jean-Marc BERNARD.

l'utilisateur. C'est précisément l'objectif d'un système d'évaluation que de mettre en cohérence les différentes utilisations de l'information et de s'assurer d'une circulation optimale de celle-ci. Pour cela, **il est indispensable de définir les responsabilités de chaque entité institutionnelle tant au niveau de la collecte de l'information, de son analyse que de la prise de décision.**

3.4.2. La disponibilité et la prise en compte des informations

Trois types d'information pourront être mobilisés ou produits selon le groupe, il s'agit :

- a. Des données issues d'**examens nationaux** (en particulier de fin de cycle primaire et de fin de cycle moyen) qui constituent une information très répandue, mais largement sous-utilisée dans beaucoup de pays. Il convient, dans le cadre d'un dispositif d'évaluation, d'en assurer la fiabilité, de la rendre disponible par établissement et de calculer aux différents échelons des valeurs moyennes et des écarts. Les établissements seront les premiers bénéficiaires de cette information, qui les aidera à se situer par rapport à d'autres et à améliorer leur performance ; les communautés éducatives en profiteront par ricochet. Les inspecteurs (ou l'échelon supérieur) s'en serviront également pour repérer les établissements en difficulté nécessitant des actions de remédiation.

Cependant, les données d'examens nationaux présentent des limites. Pour assurer un suivi plus régulier tout au long du cycle, des informations complémentaires sont nécessaires, en particulier :

- b. Des données provenant d'**épreuves communes** administrées de manière plus ciblée, par exemple à l'issue des différents sous-cycles. Il importe que ces épreuves communes réduisent la part de subjectivité locale dans la mesure des progressions par rapport aux objectifs des curricula. Ainsi, tôt ou tard, il faudra considérer le passage d'une logique d'épreuves communes à celle d'épreuves véritablement standardisées (c'est-à-dire des épreuves communes administrées dans les mêmes conditions de passation et dont la correction est centralisée).¹⁸
- c. Des données issues d'**épreuves standardisées** peuvent être recueillies, notamment, en développant au niveau national des banques d'items. En plus de mettre l'évaluation directement au service de l'atteinte des objectifs des programmes scolaires, ce type d'information rendra possible la comparaison des performances en matière d'acquisitions. C'est également sur la base de cet instrument national que pourra être envisagée une éventuelle articulation avec des instruments d'évaluation utilisés à l'échelon international.

Ces trois types de données ainsi collectées doivent motiver des actions de remédiation à conduire sur le terrain (auprès des établissements, des professeurs, etc.). Il sera ainsi valorisé la fonction de l'évaluation comme outil de soutien à l'apprentissage.

¹⁸ Des décisions opérationnelles comme des décisions de redoublement pourraient également être améliorées et rendues plus équitables de cette manière.

À noter que selon les réalités respectives des systèmes éducatifs, d'autres dimensions peuvent être prises en compte pour fin d'évaluation. Le document de réflexion et d'orientation *Pour une nouvelle dynamique de la gestion scolaire* (CONFEMEN 2007) présente en effet des éléments jugés indispensables au bon pilotage d'un système scolaire. Ainsi, la formation des enseignants, la disponibilité et la qualité des manuels scolaires, la responsabilisation et l'implication des différents partenaires pourraient constituer des éléments d'évaluation additionnels qui viendraient compléter les données des acquis des élèves. Car, faut-il le rappeler, l'évaluation doit être et doit demeurer multidimensionnelle et multipartite afin de permettre des interventions judicieuses.

Il est utile encore une fois de distinguer l'information produite dans le fonctionnement quotidien du système éducatif de l'information qui nécessite des études complémentaires. Dans le premier cas, la définition claire des responsabilités est fondamentale pour une bonne circulation de l'information. Dans le second cas de figure, c'est la question de l'attribution de ces études complémentaires à des structures du ministère qui doit faire l'objet d'une attention particulière. Il est important de conserver à l'esprit que l'évaluation n'est pas l'apanage d'une seule structure mais est l'affaire de toutes les structures. Néanmoins, l'information doit être à un moment centralisée et synthétisée afin de pouvoir fournir un tableau de bord de l'ensemble du système éducatif. Il ne s'agit pas uniquement de recueillir de l'information, il faut également mener des analyses parfois complexes demandant des compétences spécifiques. Une structure doit donc avoir clairement en charge ces différents aspects.

La fiabilité et la pertinence des informations recueillies dépendent du traitement que l'on veut opérer et de la nature des indicateurs que l'on souhaite obtenir. La définition précoce, précise et partagée entre tous les acteurs des indicateurs du tableau de bord est donc la clef de voûte de la stratégie d'information.

3.4.3. La mise en place d'une stratégie de communication

L'évaluation remplit une fonction d'information à l'égard de l'ensemble des acteurs et des partenaires de l'école. Cela suppose la définition d'une politique de communication orientée vers chacun de ces multiples destinataires. Une stratégie de partage de l'information suppose l'adaptation différenciée des messages à plusieurs groupes de destinataires, en fonction de leur intérêt. Les médias également doivent être différenciés, en fonction de leurs spécificités et des publics cibles.

C'est d'une véritable stratégie de communication qu'il faut parler ici, en prenant en compte les publics cibles, le format des informations et des messages à diffuser, les supports et les vecteurs de cette diffusion. On distinguera globalement trois types de public cible, à savoir les experts, les décideurs et les acteurs intermédiaires ou proches du terrain.

3.4.4. L'implication et la participation des acteurs aux processus d'évaluation

3.4.4.1. Les experts

Les résultats d'une évaluation se présentent le plus souvent comme de gros rapports touffus, comportant beaucoup de chiffres et de tableaux parfois compliqués, en raison de la complexité des domaines étudiés, de la technicité des méthodes employées et du degré d'élaboration des indicateurs mis en place. Il importe de garantir la pertinence des résultats

et de leur conférer la valeur la plus objective possible, en adoptant une démarche scientifique.

De tels documents sont évidemment destinés à une communauté de spécialistes, sur le plan national comme sur le plan international. Par exemple, les auteurs du rapport de synthèse du PASEC (CONFEMEN, 1999) le savent bien. Leur ouvrage peut s'aborder par une lecture rapide, non-technique, suffisante pour des non-spécialistes, tandis que des développements plus précis et plus approfondis s'adressent aux scientifiques (et aux décideurs qui voudraient connaître les fondements scientifiques des résultats).

Sous cette forme scientifiquement et techniquement élaborée, la diffusion se fait, classiquement, au travers de publications, de conférences, de participation à des séminaires ou des colloques. La mise en ligne sur site Internet permet d'utiliser les nouveaux vecteurs de communication en facilitant l'accès à l'information, en créant des liens et des référencements auprès d'autres organismes de coopération internationale.

3.4.4.2. Les décideurs

D'autres formats sont cependant nécessaires : les décideurs politiques n'ont pas les mêmes attentes que les experts en matière d'évaluation, et ce dont ils ont besoin d'abord, ce sont des recommandations fondées et fiables, qui puissent constituer une aide à la décision, et contribuer efficacement au pilotage des systèmes. C'est pourquoi il est nécessaire de reformuler de façon moins technique et plus opérationnelle les conclusions et recommandations auxquelles aboutit l'évaluation commanditée par ces mêmes décideurs. Car c'est le commanditaire qui décidera du statut — public, restreint ou confidentiel — à accorder à telle ou telle évaluation. La restitution se fait généralement en réunion de travail, au niveau ministériel ou parfois interministériel. Des plaquettes de vulgarisation des résultats peuvent constituer un support de communication adéquat.

3.4.4.3. Les acteurs du terrain

Par ailleurs, si l'on se soucie d'informer les acteurs proches du terrain des résultats de l'évaluation, on s'adresse cette fois à des agents de l'éducation et à des publics qui peuvent être peu au fait de la culture de l'évaluation. Il s'agit de tous les autres acteurs et bénéficiaires du système, telle que la société civile, dont les parents d'élèves, les apprenants, et les employeurs. Dans ce cas, il est évidemment nécessaire de **produire des documents vulgarisés et spécifiques**, susceptibles de fournir une information adéquate et recevable par ces différents publics, de manière à répondre à leurs attentes.

Cela suppose aussi que l'on s'interroge sur les **modalités de transmission**, qu'il s'agisse d'écrits, d'utilisation d'autres médias comme la télévision ou du média encore le plus répandu en dehors des zones urbaines qu'est la radio ainsi que lors de réunions ou de visites d'écoles ou de villages. L'essentiel étant de se convaincre que toute évaluation doit faire l'objet d'une véritable stratégie de communication et de dissémination.

Pour ce faire, l'évaluation doit être conçue comme un processus permanent de dialogue et de concertation entre les parties prenantes. Un tel processus permettra à chaque acteur de jouer un rôle actif dans l'opération. C'est ainsi que, dans le cadre des politiques éducatives actuellement mises en œuvre, le mécanisme d'évaluation qui est mis en place inclut

l'institution de cadres formels de concertation périodique entre les différents acteurs impliqués dans la mise en œuvre et le financement de ladite politique. Ces concertations sont ouvertes aux différentes parties représentées à chaque niveau.

3.4.5. Le réinvestissement des résultats

La question de la mission de compte-rendu, jouée par tout dispositif d'évaluation, a été abordée plus haut. En aval de cette mission, se pose la question de l'utilisation des résultats des évaluations dans les systèmes éducatifs. Sur ce terrain, l'état des lieux qui ressort du questionnaire adressé aux États et gouvernements est assez décevant. En effet, même dans des pays du Nord appartenant à la Francophonie et s'étant dotés de dispositifs d'évaluation complets, sophistiqués et coûteux, produisant des informations abondantes de façon récurrente, la question de l'utilisation des résultats apparaît quelque peu laissée au hasard. En France, par exemple, une fois les résultats diffusés (de façon à la fois variée et systématique il est vrai), c'est au débat public de saisir l'occasion, ou de ne pas la saisir, pour mettre à l'ordre du jour des décisions politiques ou le renouvellement des pratiques.

Aujourd'hui, parmi les États et gouvernements ayant participé à l'enquête, seul le gouvernement du Québec (Canada) cite des dispositions contraignantes quant aux résultats des évaluations : *«le sous-ministre est responsable du suivi de l'évaluation et s'assure que les décisions découlant des résultats mènent à l'action»*¹⁹.

Amener les décideurs à réinvestir les résultats suppose un effort spécifique en matière de communication. En particulier, on peut penser qu'un large partage des informations sur les résultats en direction de la société civile peut amener cette dernière à jouer un rôle moteur essentiel dans ce réinvestissement.

Comme indiqué plus haut, l'évaluation remplit au moins deux fonctions principales : elle vise à faciliter la prise de décision, mais elle a aussi un rôle fondamental de compte-rendu à la société, dans la mesure où il s'agit largement de fonds publics, dont l'utilisation, en démocratie, concerne donc chaque citoyen.

Au delà de ces visées générales, se profile un enjeu considérable. En effet, si l'éducation doit en premier lieu bénéficier aux élèves, son développement et sa qualité concernent également et, fort directement, de nombreuses catégories d'acteurs, à savoir :

- non seulement le ministre en charge de l'Éducation, mais l'ensemble du gouvernement, s'agissant d'un secteur particulièrement important à tous égards ;
- le ministre en charge de l'Éducation, et l'ensemble de son administration, aux niveaux central et régional, aux niveaux des écoles et des personnels de direction et d'enseignement ;
- les bailleurs de fonds internationaux et les représentants des coopérations bilatérales ;
- la société civile dans son ensemble, qu'il s'agisse des parents, des associations de parents d'élèves, des syndicats, des responsables de communautés villageoises, de

¹⁹ Réponse obtenue de la part du gouvernement du Québec (Canada) à la suite de l'enquête réalisée par et auprès des pays membres de la CONFEMEN sur l'état des lieux des dispositifs d'évaluation des systèmes éducatifs nationaux.

communautés religieuses, des organisations non gouvernementales intervenant dans le secteur ou des acteurs privés parfois sollicités pour intervenir dans le financement du dispositif.

Il est indispensable, pour que la fonction de compte-rendu de l'évaluation soit remplie, que l'ensemble de ces acteurs puisse accéder, de façon transparente et adaptée, aux informations sur l'état du système éducatif dans le pays.

En effet, si les responsables politiques, les bailleurs et les experts disposent d'un accès privilégié aux informations, y compris les plus sensibles, il n'en va pas de même pour les autres acteurs. On retiendra ici deux exemples :

■ **Les informations collectées** lors de la réalisation d'une évaluation proviennent le plus souvent des acteurs intermédiaires ou de terrain. Elles sont obtenues, par exemple, à partir de statistiques, de questionnaires, parfois d'entretiens (directeurs d'école, parents, etc.). Or, si ces informations (souvent accompagnées de requêtes), «remontent» bien vers les décideurs, il est plus rare que leurs résultats «redescendent» vers les acteurs qui ont pourtant fourni nombre des éléments indispensables. Il en découle parfois une amertume, voire une défiance, qui n'incite pas toujours à renouveler la participation ou qui, si elle est une obligation de service pour tel ou tel agent, l'incite à y répondre de façon bureaucratique et, en quelque sorte, «ritualisée». Il en résulte ainsi un sentiment de déresponsabilisation qui peut induire nombre d'effets pervers sur la qualité de l'évaluation elle-même.

■ **Le «rôle des familles»** est souvent souligné comme étant un facteur décisif de l'amélioration de la qualité de l'éducation. Or, les familles et les communautés sont souvent incitées à participer au coût de l'éducation. Cela concerne dans presque tous les cas des frais de scolarisation. Cela peut concerner également l'habillement, la nourriture, la fourniture de livres, voire la construction de classes dans certains cas. De telles pratiques vont bien sûr à l'encontre de l'objectif de gratuité de l'éducation primaire ou de base, mais la contrainte financière est souvent telle qu'elles sont jugées nécessaires, au moins pour le moment dans certains pays.

Dans ces conditions, ne pas fournir d'information sur les résultats obtenus, sur le sens de l'effort demandé (parfois perçu comme excessif) risque de provoquer des effets très négatifs, car d'une façon ou d'une autre, chacun se forme une opinion sur l'éducation dispensée à ses enfants. La laisser se former à partir d'impressions ou d'intuitions erronées ou mal informées peut être dangereux. Au total, ne pas entretenir un dialogue avec ces acteurs concernés au premier chef peut avoir de graves effets de démotivation, provoquer du mécontentement et des abandons de la scolarisation, et influencer négativement sur la demande d'éducation.

Or, il est clair que dans certains pays, ou dans certaines zones, l'un des problèmes les plus importants à résoudre est celui de l'ajustement de l'offre et de la demande d'éducation. On sait qu'un système éducatif ne peut être entièrement régulé par l'offre d'éducation (à supposer même qu'elle soit toujours suffisante). On sait aussi que parfois, il faut stimuler et entretenir la demande d'éducation.

Dans cette recherche, d'un point d'ajustement réaliste entre offre et demande d'éducation, les informations issues de l'activité d'évaluation, une fois mises à disposition sous une forme et des modalités adaptées, peuvent apporter une importante contribution.

CHAPITRE IV :

LES PERSPECTIVES D’ACTION

4.1. LES RÉSULTATS ATTENDUS DE L'INSTITUTIONNALISATION DE L'ÉVALUATION

L'évaluation doit être conduite dans le but d'améliorer la mise en œuvre des politiques éducatives adaptées aux réalités nationales et locales, afin d'assurer une gestion efficace et efficiente des systèmes éducatifs.

4.1.1. ...au plan du pilotage pédagogique

L'évaluation doit permettre d'identifier les modes d'organisation les plus pertinents ou les plus efficaces dans un contexte local, national, ou régional donné. Pour ce faire, il s'agira d'identifier les facteurs scolaires (manuels scolaires, matériel didactique, méthodes et temps d'enseignement, curricula, relations pédagogiques, formation des enseignants) les plus pertinents pour un enseignement de qualité. L'évaluation doit également donner des indications sur les succès ou les échecs des réformes pédagogiques (programmes, méthodes d'enseignement, manuels scolaires, formations initiale et continue, etc.) et proposer des mesures de remédiation.

4.1.2. ...au plan du pilotage administratif

L'évaluation constitue un instrument d'identification des forces et des faiblesses des dispositifs de pilotage des systèmes éducatifs (les dispositifs administratifs de gestion aux niveaux central, régional et local). Elle permet de promouvoir de nouvelles modalités de gestion à tous les niveaux du système, fondées sur une démarche de pilotage par les résultats dans le cadre d'une dynamique partenariale engageant l'ensemble de la communauté éducative. Elle permet également d'accroître la transparence, la responsabilité et l'imputabilité pour une meilleure gouvernance.

4.1.3. ...au plan économique et financier

Le secteur éducatif, auquel sont fréquemment consacrés des montants considérables de ressources, (de 3 % à 6 % du produit intérieur brut (PIB), souvent plus du cinquième des ressources publiques), et dont chacun attend beaucoup, tant au niveau individuel que collectif, doit intégrer ce mouvement global de développement de l'évaluation des politiques publiques. On ne peut passer sous silence que l'investissement pour évaluer a un coût, notamment un coût financier. En effet, la mobilisation de ressources conséquentes à même de garantir la qualité des évaluations et leur nombre permet d'assurer une meilleure mesure pour les politiques.

Du point de vue de la reddition des comptes, l'enjeu premier de l'évaluation des politiques éducatives est celui de l'efficacité, il s'agit de s'assurer qu'il existe un rapport convenable entre les coûts et les résultats du système éducatif.

Concernant la recherche de réduction des coûts unitaires de l'éducation, il est évident que les mesures prises ont ou auront un impact direct sur les populations, car elles sont immédiatement visibles, qu'il s'agisse du niveau de recrutement et de rémunération des maîtres, de la taille des classes, des redoublements, etc. Il faut donc pouvoir expliquer et justifier de telles mesures, notamment en montrant qu'elles ne se font pas au détriment de la qualité de l'éducation dispensée.

De même, la recherche de ressources additionnelles pour le secteur amène souvent les décideurs, malgré leur adhésion à l'objectif de gratuité de l'éducation de base, à rechercher du côté des communautés, voire de la société civile, des financements complémentaires pour réduire les contraintes financières auxquelles ils font face.

En effet, les financements privés (au sens strict du terme) interviennent fort peu dans les systèmes éducatifs, et lorsqu'ils le font, c'est bien souvent au niveau de l'enseignement secondaire et surtout supérieur qu'au niveau de l'éducation de base. Cependant, ici ou là, quelques entreprises et quelques banques ont développé des actions ponctuelles qui conduisent à une collecte de fonds transitant le plus souvent par le canal d'ONG. D'autres fonds privés peuvent être d'origine confessionnelle ou encore provenir de fondations.

Mais la principale ressource complémentaire pour l'éducation de base provient des frais de scolarité demandés aux familles, de leurs contributions directes au fonctionnement de l'école (paiement des salaires des enseignants, achats de matériels, construction des écoles, etc.), auxquels s'ajoutent l'habillement, la nourriture, l'achat de livres pour les élèves, etc. Enfin, dans bon nombre de pays, les écoles développent des activités génératrices de revenus, comme l'élevage, la culture potagère, la couture, etc.

L'effort demandé aux parents et aux communautés est donc une démarche délicate, en ceci qu'elle exige la compréhension et la collaboration des intéressés. Or, des charges perçues comme trop lourdes risquent d'être mal supportées et de provoquer des mécontentements ainsi que des mouvements de déscolarisation, notamment dans les milieux les plus fragiles économiquement et socialement. Toutefois, bien que les coûts soient un facteur qui conduit au retrait de la vie scolaire de plusieurs enfants, il faut noter que le jugement défavorable sur l'utilité de l'école constitue une contrainte majeure.

Les collectivités territoriales, dans la plupart des pays, participent aux financements des systèmes éducatifs sans toujours les articuler au programme de développement de l'éducation. L'évaluation permettra de mieux réorienter leurs interventions vers des actions en harmonie avec le programme sectoriel.

Les aspects financiers de la gestion des systèmes éducatifs peuvent également bénéficier des évaluations, en ce qu'elles mettent en relation les résultats obtenus avec les moyens mobilisés. Elles indiquent les domaines et actions les plus et les moins efficaces.

4.1.4. ...au plan social et politique

La gouvernance constitue le substrat sur lequel repose la gestion du développement ; elle en est un des «dynamiseurs», car les faiblesses en matière de gouvernance affectent l'économie dans son ensemble. De même, un développement économique et social durable influe sur les conditions d'instauration de la démocratie. Aussi, dans l'approche du développement, le principe d'une bonne gouvernance²⁰ s'impose. C'est dire que l'évaluation des politiques publiques de développement passe aussi par celle de la gouvernance dont les éléments clés sont, faut-il le rappeler, la responsabilité, l'État de droit et la transparence. Actuellement, la mise en œuvre des politiques de développement (macro-économiques et

²⁰ La «bonne gouvernance» s'entend d'un système démocratique de gestion des affaires d'intérêt général dont les caractéristiques principales sont : la responsabilité, la soumission de l'État au droit, la transparence et l'honnêteté.

sectorielles) repose sur un certain nombre de principes directeurs : **approche participative, décentralisation, intégration** des considérations **du genre** et des aspects environnementaux, **transparence** dans la gestion, et mise en place d'une **stratégie d'information et de communication**.

De plus, l'évaluation est un des outils permettant à la société de se situer, de comprendre son propre fonctionnement tout en prenant conscience de son état. L'évaluation permet ainsi de développer la culture du compte-rendu permettant à la société d'augmenter ses attentes, ses exigences et ses demandes sociales en éducation. Dans cette optique, la société joue un rôle actif en étant à la fois bénéficiaire et commanditaire de l'évaluation et de ses impacts sur le système éducatif. La société civile (syndicats, associations, ONG, etc.) et la communauté reconnaissent de plus en plus leur rôle de commanditaire qui est en droit de demander qu'il soit rendu compte de l'intervention publique et de l'utilisation des ressources. La demande sociale d'éducation ne s'exprime pas en effet sur des marchés, pour l'essentiel ; on sait cependant que le développement du secteur n'est pas le simple fait d'une offre publique accrue en éducation. L'évaluation des politiques éducatives, en ce qu'elle fournit une occasion d'associer de multiples partenaires, permet à tous de partager des objectifs sociaux ainsi que de s'exprimer sur les moyens à mettre en œuvre pour les atteindre. Elle permet également aux autorités publiques de faire émerger un point de vue distinct de celui exprimé par les administrations.

Une des finalités de l'éducation est de développer les ressources humaines dont le pays a besoin pour assurer son développement économique et social durable, et ce, tout en tenant compte des besoins du marché en perpétuelle évolution. Ce qui se mesure par le degré d'insertion socioprofessionnelle des sortants du système éducatif (efficacité socio-économique ou efficacité externe du système éducatif). Or, dans bon nombre de pays, notamment ceux en développement, les «produits de l'entreprise école» correspondent peu ou pas aux besoins de l'économie.

Aussi, il est indispensable d'intégrer dans la définition des politiques éducatives, une dimension sociologique qui prenne en compte à la fois les situations environnementales, économiques et culturelles. Pour cela, une démarche spécifique d'évaluation, construite en fonction de ces aspects, est de nature à contribuer à mieux cerner les différents facteurs intervenant dans la formulation de la demande d'éducation, à orienter des choix et, ainsi, à concourir à un meilleur ajustement entre l'offre et la demande d'éducation, à des réponses publiques en termes d'offre qui correspondent le plus étroitement possible aux besoins de l'économie nationale et de la société.

Au total, il apparaît clairement que l'évaluation peut et devrait sans doute occuper une place centrale dans les procédures de pilotage des systèmes éducatifs, tant pour des raisons externes (engagements internationaux) qu'internes (nécessité de piloter et de réguler les politiques et stratégies éducatives, nécessité de rendre compte, nécessité de rationaliser l'utilisation des ressources, etc.). Dans cette perspective, il y aurait des avantages certains à créer un dispositif d'évaluation intégré à la vie du système ou, là où il existe déjà, à le renforcer ou à l'intégrer au mieux aux procédures de régulation de l'ensemble du système éducatif.

Enfin, chacun observe une lente évolution de l'ensemble des systèmes éducatifs, qui est marquée par une logique souvent décrite de convergence des formes. Les différences les

plus saillantes d'un système national à l'autre tendent à s'estomper, qu'il s'agisse du schéma général des études (exemple : le système LMD (Licence, master, doctorat) qui se généralise), des curricula, des méthodes ou des contenus (tous fortement influencés par l'APC (Approche par les compétences)). Dans ce contexte, l'évaluation des systèmes et de leurs performances prend un relief particulier en ceci qu'elle permet, à intervalles réguliers, de situer un système national en comparaison de systèmes voisins.

Enfin, à un niveau plus global de régulation politique et sociale, les résultats des évaluations sont attendus pour nourrir et structurer les débats qui se nouent autour de questions scolaires.

4.2. DES CONDITIONS POUR L'INSTITUTIONNALISATION DE L'ÉVALUATION

4.2.1. Une volonté politique réelle et effective

L'existence d'une volonté politique réelle, le développement d'une culture d'évaluation auprès de tous les acteurs du système éducatif ainsi que la mise en place et la formation des structures nationales constituent les préalables au développement d'un système national d'évaluation.

La valorisation des systèmes nationaux d'évaluation et leur articulation par rapport aux initiatives et dispositifs internationaux passent par la contextualisation des outils d'évaluation, la comparaison internationale des résultats pour situer les spécificités nationales et l'identification d'objectifs définis par les besoins nationaux.

Soulignons que l'évaluation doit être **transparente et indépendante**. Il est indispensable d'assurer la transparence à tous les niveaux de la démarche évaluative, qu'il s'agisse des objectifs, de la collecte des données, du croisement de celles-ci avec de multiples interlocuteurs, de la diffusion des résultats, etc. C'est à cette condition *sine qua non* qu'une évaluation peut se donner les chances de parvenir à des résultats de qualité. La recherche de la transparence peut et doit permettre la mise en place d'une relation de confiance avec les interlocuteurs. Mais elle doit également se prévaloir d'une mission claire qui lui donne autorité pour conduire ses travaux. S'agissant d'évaluation d'organes centraux du ministère, et compte tenu de son caractère sensible, cette autorité ne peut lui venir que du commanditaire, c'est-à-dire le ministre lui-même.

L'indépendance représente le corollaire direct de la transparence. En effet, c'est de la mission confiée par l'autorité supérieure que les évaluateurs tirent leur légitimité. Mais c'est aussi par leur indépendance vis-à-vis des agents et des responsables des organismes évalués qu'ils peuvent chercher à obtenir transparence et coopération. Enfin, à l'égard du commanditaire, les évaluateurs bénéficient également d'indépendance : si un cahier des charges leur est proposé, celui-ci ne dit rien quant à la nature et au contenu des résultats attendus. Il appartient au commanditaire de faire siennes ou de rejeter les conclusions et recommandations auxquelles parvient l'évaluation, sans préjudice pour les évaluateurs qui n'ont qu'une obligation, celle de la rigueur méthodologique pour garantir la qualité de leur démarche et de leurs résultats.

4.2.2. La mise en place de dispositifs d'évaluation

Avant de développer un dispositif national d'évaluation, il faut d'abord le créer et lui assigner des objectifs précis, spécifiques à chaque pays.

Le questionnaire réalisé sous l'égide de la CONFEMEN a montré que la plupart de ses pays membres ne dispose pas aujourd'hui d'un véritable dispositif national d'évaluation centré sur la mesure des performances des établissements, même si la majorité des pays a participé à des opérations d'évaluation des acquis des élèves et même si ces évaluations comportent systématiquement des objectifs de transfert : encadrement d'équipes nationales qui sont censées maîtriser les techniques utilisées, constitution d'embryons de systèmes nationaux d'évaluation.

Il en va de même des recherches spécifiques associées à la participation du pays à un programme international d'évaluation des systèmes éducatifs : ces recherches se basent toujours sur des échantillons et ne permettent pas une gestion courante de l'ensemble des établissements d'un pays ; elles mettent le plus souvent en œuvre une méthodologie et un cadre d'analyse complexes qui ne favorisent pas toujours le développement d'une culture nationale d'évaluation.

La conception d'un dispositif national d'évaluation doit être clairement distinguée dans ses objectifs et dans ses contours des évaluations ponctuelles ou des recherches spécifiques associées à la participation du pays à un programme d'évaluation des systèmes éducatifs. Toutefois, il va de soi qu'à terme un dispositif national d'évaluation doit pouvoir prendre en charge ce type d'évaluations et de recherches.

Pour la CONFEMEN, l'objectif essentiel d'un dispositif national d'évaluation devrait être d'améliorer la qualité du pilotage d'un système éducatif par les acteurs qui en ont en priorité la charge, c'est-à-dire les planificateurs, les corps d'inspection, les responsables d'établissements et les équipes éducatives.

Plus que des opérations ponctuelles d'évaluation de grande envergure pilotée plus ou moins de l'extérieur et/ou la formation de quelques spécialistes de très haut niveau, la construction d'un dispositif national d'évaluation exige d'abord un enracinement dans le fonctionnement courant des systèmes éducatifs afin de faciliter l'émergence d'une véritable culture d'évaluation parmi ses différents acteurs, qu'ils oeuvrent au plan national, régional ou local.

Deux préalables peuvent être mis en évidence pour ce faire :

- la transparence des résultats et des ressources qui permet à tous les acteurs d'évaluer leur situation et de se comparer aux autres établissements;
- la mise en œuvre effective du contrôle du fonctionnement des établissements par les acteurs qui en ont la charge car c'est ce contrôle qui permet d'agir sur les facteurs de la qualité.

Sur cette base, on peut concevoir un dispositif national d'évaluation sur le mode d'une structure nationale stable composée de personnels formés et impliquant les acteurs aux différents niveaux du système éducatif.

Cette structure pourrait avoir pour missions :

- d'énoncer les motifs et avantages de l'évaluation aux niveaux régional et local dans le but d'une meilleure implication des acteurs dans le processus ;
- de responsabiliser les niveaux régional et local dans les prises de décision relativement aux recommandations découlant des évaluations ;
- d'assurer une articulation entre la structure nationale d'évaluation et celles responsables des évaluations aux niveaux régional et local.

Cette appropriation progressive des enjeux et des méthodes d'évaluation constitue le socle de la diversification et de la complexification du dispositif national d'évaluation, qui peut ensuite intégrer des équipes de spécialistes aptes à conduire des évaluations plus ambitieuses mais aussi participer au plan national à l'expérimentation et à la validation de nouvelles méthodes et formes d'organisation pédagogiques. Elle devrait, selon la CONFEMEN, s'appuyer sur le levier que constitue l'établissement scolaire, susceptible de permettre une régulation des comportements et des activités.

4.2.3. La mise à disposition des ressources appropriées

La conduite d'une évaluation, processus complexe et multidimensionnel, nécessite des ressources conséquentes. Pour que les dispositifs nationaux puissent répondre aux questions précises posées par les évaluations (nationales et internationales) et surtout trouver des pistes d'amélioration, il est nécessaire non seulement que ceux-ci soient pensés en partant de l'échelon local, mais aussi qu'ils soient pourvus en ressources appropriées.

En général, pour les pays du Nord, ces ressources nécessaires à la conduite des évaluations (ressources organisationnelles, humaines, instrumentales, logistiques et financières) sont disponibles. En revanche, pour les pays du Sud, aux moyens modestes et où la pratique de l'évaluation n'est pas encore totalement ancrée dans les mœurs de gestion des systèmes éducatifs, ces ressources font souvent défaut. Il convient cependant de souligner que la coopération internationale supplée le plus souvent à ce déficit, à travers des appuis visant le renforcement des capacités des ressources humaines et logistiques. Même dans ce cas de figure, il arrive parfois que les ressources budgétaires des États qui devraient financer le fonctionnement ordinaire des dispositifs nationaux d'évaluation, ne puissent pas être mises à la disposition des services chargés de conduire les actions d'évaluation.

Ainsi, l'enjeu pour plusieurs États et gouvernements francophones, notamment ceux du Sud, est de créer, à partir d'une situation d'information et de répartition des rôles, de compétences humaines et ressources financières insuffisantes, un véritable système d'évaluation centré sur les enjeux premiers de scolarisation. Ce qui nécessite l'appui de la coopération internationale.

4.3. L'ARTICULATION ENTRE LES ACTIVITÉS NATIONALES, RÉGIONALES ET INTERNATIONALES

Aujourd'hui, les activités d'évaluation menées au sein de chaque système national sont souvent peu articulées aux programmes internationaux d'évaluation : les structures chargées de la mise en oeuvre diffèrent le plus souvent, ainsi que les outils et les modes de diffusion des résultats.

Il ne devrait pas y avoir de confusion entre les programmes internationaux (PASEC, SACMEQ, etc.) et l'appui qu'il conviendra d'apporter aux structures nationales d'évaluation. Il ne s'agit pas de reproduire des évaluations de type PASEC ou SACMEQ au niveau national, mais il y a la nécessité d'une approche harmonisée du fait que les pays sont presque tous quasiment confrontés à différentes difficultés. Certes, le dispositif national d'évaluation ne peut pas fonctionner en autarcie, mais son ouverture vers le niveau international doit être coordonnée.

L'articulation des dispositifs nationaux et internationaux ainsi que la possibilité d'établir des comparaisons internationales s'avèrent de plus en plus nécessaires. Cependant, c'est au niveau de chaque pays que l'initiative est prise de mettre en œuvre des dispositifs d'évaluation.

Le rôle essentiel des structures d'évaluation internationales consiste en l'appui à la mise en place et au renforcement du dispositif national d'évaluation. La nature de cet appui dont l'objectif fondamental est de permettre aux structures d'évaluation d'atteindre leurs objectifs se décline en quatre axes :

- **Renforcement des capacités nationales** : les acteurs qui composent les structures d'évaluation nationales doivent être formés au développement des instruments, aux techniques d'analyse, etc.
- **Appui financier** : les institutions de financement internationales devraient appuyer la mise en place des dispositifs nationaux.
- **Accompagnement scientifique et technique** : des programmes internationaux qui ont fait leur preuve dans le domaine de l'évaluation, tels que le PASEC, le SACMEQ ou le PISA, devraient appuyer les dispositifs nationaux sur les plans scientifique et technique. Il est souhaitable que le PASEC agisse comme pôle d'expertise, en assurant la veille scientifique et technique sur les outils et méthodes d'évaluation et en proposant leur adaptation aux contextes régionaux. Ce programme d'évaluation veillera, conformément à son mandat original, à diffuser les outils et méthodes d'évaluation et à la création interne et permanente d'évaluation des systèmes éducatifs.
- **Coordination des échanges d'expériences et de pratiques** : les structures internationales devraient favoriser les échanges entre les différentes structures nationales et organiser des cadres d'échanges, notamment sur les expériences et les bonnes pratiques qui se développent au niveau des différents échelons nationaux. À titre d'exemple, une plateforme virtuelle pourrait être créée pour supporter ces initiatives.

C'est à la structure nationale d'évaluation que reviendrait le rôle d'assurer l'appropriation du dispositif par les acteurs, la sensibilisation aux enjeux et pratiques du pilotage par les résultats au sein des systèmes éducatifs, et d'assumer la responsabilité de sa mise en œuvre. Le déroulement des étapes subséquentes de la mise en place du dispositif dépendrait par conséquent des ressources dont il dispose : il devra bénéficier des moyens adaptés à sa mission, tant sur le plan des ressources humaines (identification des profils des membres et formation) que des ressources matérielles et financières.

Pour aider la structure nationale à cerner les éventuelles difficultés et à ajuster les actions, des activités de suivi et d'évaluation devraient être prévues. Par ailleurs, dans la mesure où des actions de nature comparable seraient conduites dans plusieurs pays, il devrait être envisagé d'organiser des activités de partage d'expériences.

4.4. LE RÔLE DE LA CONFEMEN

La CONFEMEN est la plus ancienne institution de la Francophonie. Depuis sa création en 1960, la CONFEMEN oeuvre pour la promotion de l'éducation et de la formation professionnelle et technique. Elle compte aujourd'hui 41 États et gouvernements membres.

Elle constitue l'instrument privilégié qui permet aux États et gouvernements, tant du Nord que du Sud, de conjuguer leurs efforts en vue de développer leurs systèmes éducatifs. Son rôle est d'orienter les politiques éducatives dans les pays ayant le français en partage, à travers les décisions et recommandations prises lors des sessions ministérielles.

La CONFEMEN dispose d'une expertise déjà longue et reconnue en matière d'évaluation, notamment au travers de son Programme d'analyse des systèmes éducatifs (PASEC), créé en 1991 dans la foulée de la première Conférence mondiale sur l'Éducation pour tous de Jomtien (1990). Cet outil d'analyse et d'évaluation fournit aux décideurs des informations objectives pour orienter leurs choix en matière de politiques éducatives.

Le programme a généré de nombreuses études et publications, études diagnostiques ou thématiques, qui ont permis de nombreux apports dans la réalisation de son premier objectif. De plus, par ses nombreuses actions de renforcement de capacités (missions, formations en ateliers, supports méthodologiques, etc.), il contribue au développement de la culture d'évaluation et à la diffusion de ses méthodes.

La CONFEMEN doit occuper toute sa place dans ce débat essentiel en s'appuyant, notamment, sur le programme PASEC qui est déjà une expérience dans ce domaine et qui lui confère une forte légitimité. L'enjeu est de faire émerger des solutions concrètes sous la forme de dispositifs nationaux d'évaluation, mis en œuvre par les États membres de la CONFEMEN avec l'appui de l'OIF et approuvés par les partenaires techniques et financiers.

La CONFEMEN aura plusieurs rôles à jouer dans cet exercice :

- définir et diffuser des orientations sur l'évaluation des systèmes éducatifs ;
- promouvoir les échanges d'expériences et de bonnes pratiques ;
- assurer la coordination et l'harmonisation des politiques d'évaluation adoptées par les États et gouvernements membres ;
- accompagner scientifiquement les projets des pays membres ;
- promouvoir les collaborations entre pays dans une perspective comparative ;
- faire un plaidoyer au plan international et auprès des pays membres sur l'importance de l'évaluation dans le pilotage des systèmes éducatifs.

L'objectif général visé par toutes ces actions est de permettre aux pays d'évaluer et de suivre la mise en œuvre de leurs politiques éducatives afin de pouvoir prendre à chaque niveau les décisions d'amélioration qui s'imposent, de contribuer à un pilotage par les résultats de leurs systèmes éducatifs et d'entrer progressivement dans une culture de redevabilité.

CONCLUSION

L'atteinte des objectifs de réduction de la pauvreté et du développement humain durable, ainsi que ceux d'universalisation de l'éducation primaire (OMD 2) exigent, entre autres, de disposer d'un système opérationnel et efficace d'évaluation. En effet, l'évaluation est essentielle pour maintenir et renforcer la volonté politique, identifier les besoins, mobiliser les ressources nécessaires et conduire des actions efficaces.

Un bon système d'évaluation dépend grandement sur l'engagement des pouvoirs publics. En effet, son succès repose en grande partie de l'attitude des décideurs et gestionnaires du système éducatif qui doivent œuvrer à répandre et à développer la culture de l'évaluation, notamment dans la manière dont sont prises leurs décisions. Par exemple, les responsables au plus haut niveau des ministères en charge de l'éducation devront s'engager et souligner l'importance de prendre des décisions sur la base de données (produites et validées dans le cadre d'un tel processus) et en s'appuyant sur des analyses sérieuses. Ceci contribuera de manière importante à asseoir une gestion transparente axée sur les résultats.

La mise en place d'un système d'évaluation au sein d'un système éducatif est une entreprise complexe qui, au-delà des ressources budgétaires devant lui être allouées, implique un certain nombre d'étapes. Tout d'abord, la mise en place d'un dispositif de ce genre n'a de sens que si l'on s'inscrit résolument dans une logique de pilotage par les résultats, intégrant non seulement les performances des élèves, mais également les évaluations des enseignants et des manuels, deux facteurs déterminants de la réussite scolaire. Il faut donc engager une réflexion sur les principaux axes d'évaluation à retenir sur la base des besoins en matière de pilotage du système. Une fois ces axes définis, il faut ensuite identifier tous les indicateurs qui permettront d'évaluer les performances du système éducatif. Il devient enfin nécessaire de confronter les besoins en information du processus d'évaluation avec les informations disponibles au sein du système éducatif. Des études peuvent être nécessaires pour compléter les informations disponibles.

Pour qu'un système d'évaluation puisse s'intégrer dans le dispositif de pilotage du système éducatif, il doit avoir nécessairement les caractéristiques suivantes :

1. porter à la fois sur les résultats, pour rendre compte et servir à l'impulsion, et sur le fonctionnement pour nourrir le suivi, l'animation, le conseil, la formation, et ainsi aider au progrès ;
2. être «participative» pour augmenter ses chances d'influer sur les décisions prises par les agents ;
3. interpellier à la fois les décideurs et tous les autres acteurs, même s'il existe diverses sortes de résultats pour permettre de mieux les atteindre dans leur diversité et s'inscrire ainsi dans une logique de pilotage concertée.

Cependant, le fonctionnement optimal des dispositifs d'évaluation exige une claire définition des responsabilités au niveau institutionnel. Un système d'évaluation ne doit pas être en marge des structures administratives existantes mais intégré à celles-ci, c'est pourquoi la dimension institutionnelle est essentielle dans ce processus. Des mécanismes de financement permettant de pérenniser les activités d'évaluation doivent être définis et mis en place (ceci devrait constituer un préalable à l'appui international), les capacités des ressources humaines développées et les mécanismes de diffusion des résultats établis.

Enfin, le système d'évaluation pour la gestion de l'éducation, pour être efficace, a besoin d'un soutien clairement exprimé de la part du (ou des) ministre(s) en charge du secteur.

BIBLIOGRAPHIE

BANQUE MONDIALE (2001) — *Une chance pour apprendre. Le savoir et le financement pour l'éducation en Afrique sub-saharienne*. Département Développement humain, Région Afrique. Banque mondiale.

BERNARD, J.M. et MICHAELOWA, K. (2006) — «Managing the impact of PASEC projects in francophone sub-Saharan African» in IIPE *Cross National Studies of the Quality of Education. Planning their design and managing their impact*. Paris: UNESCO/IIPE.

BOTTANI, N. et GAUVREAU, C. (2002) — *Les nouvelles tendances en évaluation dans un contexte d'efficacité, d'équité et de recevabilité*. Proposition au Séminaire de formation SRED.

BOTTANI N. et VRIGNAUD P. (2005) — *La France et les évaluations internationales*. Rapport établi à la demande du Haut Conseil de l'évaluation de l'école.

COLOMBO G. (2000) — «Entre système modulaire et observation continue. La problématique en Italie» dans CIEP *Revue internationale d'éducation Sèvres*, n° 26.

CONFEMEN (1999) — *Les facteurs de l'efficacité dans l'enseignement primaire : les résultats du programme PASEC sur neuf pays d'Afrique et de l'Océan indien*. Dakar : CONFEMEN.

CONFEMEN (2004) — *L'évaluation des systèmes éducatifs*. Document de travail sur le thème de la 51^e session ministérielle de Maurice.

CONFEMEN (2006) — *Actes de la 52^e session ministérielle de la CONFEMEN de Niamey (Niger)*. Dakar : CONFEMEN.

CONFEMEN (2007) — *Pour une nouvelle dynamique de la gestion scolaire*. Document de réflexion et d'orientation. Dakar : CONFEMEN.

DEMAILLY L., VAN ZANTEN A., THELOT C. et al. (2002) — «Évaluer oui mais pour quoi faire ? : dossier». Institut de recherche de la FSU, *Nouveaux regards*, n° 16, p. 2-22.

DEMAILLY L. (2000) — *Évaluer les politiques éducatives : sens, enjeux, pratiques*. Bruxelles: De Boeck Université.

DURU-BELLAT M. et JAROUSSE J.-P. (2001) — *Portée et limites d'une évaluation des politiques et des pratiques éducatives par les résultats*. IREDU, Université de Bourgogne, France.

GAUTHIER P.-L., (2000) — «Du bon usage de l'évaluation» dans CIEP *Revue internationale d'éducation Sèvres*, n° 26.

GÉRARD F.-M. (2001) — «L'évaluation de la qualité des systèmes de Formation», *Mesure et évaluation en éducation*, vol 24.

IIPE (2006) — *Cross National Studies of the quality of education, Planning their design and managing their impact*. Paris: UNESCO/IIPE.

HARDOUIN J.C, HUSSENET A., SEPTOURS G. et BOTTANI N. (2003) — *Éléments pour un diagnostic sur l'école*. Document général préparatoire au débat national sur l'avenir de l'école.

KELLAGHAN T. et GREANEY V. (2002) — *L'évaluation pour améliorer la qualité de l'enseignement*. Paris : IIEP/UNESCO.

MICHEL A. (2000) — «Évaluer pour piloter. Réflexion sur le cas français» dans CIEP *Revue internationale d'éducation* Sèvres, n° 26.

OCDE (1996) — *Évaluer et réformer les systèmes éducatifs*. Paris : OCDE.

ORIVEL F. (2003) — *De l'intérêt des comparaisons internationales en éducation*. Colloque International Education comparée et coopération, La Réunion Novembre.

PAIR C. (2001) — *Forces et faiblesses de l'évaluation du système éducatif en France*. Haut Conseil de l'évaluation de l'école. Octobre 2001.

PERRENOUD P. (1998) — *Évaluer les réformes scolaires, est-ce bien raisonnable ?* Faculté de psychologie et des sciences de l'éducation, Université de Genève.

SUCHAUT B. (2005) — *L'évaluation des dispositifs éducatifs*. Table ronde au Colloque ADMEE.

SUCHAUT B. (2006) — *Les évaluations nationales en France : Outils pédagogiques et outils de pilotage, intérêts, limites et nouvelles perspectives*. Colloque ADMEE (septembre 2006).

THELOT C. (1993) — *L'évaluation du système éducatif*. Paris : Nathan.

UNESCO (1998) — «L'évaluation des systèmes d'éducation : l'état de la question à la fin des années 90». *Perspectives*, n° 105.

UNESCO (2000) — *Cadre d'action de Dakar. L'Éducation pour tous : tenir nos engagements collectifs*. Forum mondial sur l'éducation. Dakar, 26-28 avril.

UNESCO/BREDA (2006) — *Éducation pour tous en Afrique 2006. Statistiques et analyses sous-régionales. Rapport* Dakar +6. Dakar : UNESCO/BREDA - Pôle de Dakar.