



Association pour le développement de l'éducation en Afrique

**Biennale de l'ADEA 2003
(Grand Baie, Maurice, 3-6 décembre 2003)**

**Pilotage des performances :
évaluations et examens en Afrique**

par Thomas Kellaghan et Vincent Greaney

**Document de travail
en cours d'élaboration**

NE PAS DIFFUSER

Doc. 2.D

Ce document a été commandé par l'ADEA pour sa biennale (Grand Baie, Maurice, 3-6 décembre 2003). Les points de vue et les opinions exprimés dans ce document sont ceux des auteurs et ne doivent pas être attribués à l'ADEA, à ses membres, aux organisations qui lui sont affiliées ou à toute personne agissant au nom de l'ADEA.

Le document est un document de travail en cours d'élaboration. Il a été préparé pour servir de base aux discussions de la biennale de l'ADEA et ne doit en aucun cas être diffusé dans son état actuel et à d'autres fins.

© Association pour le développement de l'éducation en Afrique (ADEA) – 2003

Association pour le développement de l'éducation en Afrique (ADEA)

Institut international de planification de l'éducation

7-9 rue Eugène Delacroix

75116 Paris, France

Tél. : +33(0)1 45 03 77 57

Fax : +33(0)1 45 03 39 65

adea@iiep.unesco.org

Site web : www.ADEAnet.org

Table des matières

ABREGE.....	9
1. RESUME.....	10
1.1. EXAMENS NATIONAUX (EXTERNES).....	10
1.2. EVALUATIONS NATIONALES	11
1.3. EVALUATIONS INTERNATIONALES	12
1.4. EVALUATIONS A L'ECHELON DE LA SALLE DE CLASSE	13
1.4.1. Problèmes liés à l'utilisation des informations tirées des évaluations à l'échelon de la salle de classe.....	13
1.5. MECANISMES D'AIDE A L'UTILISATION DES EVALUATIONS.....	14
2. CONTEXTE DES PROPOSITIONS RELATIVES A L'UTILISATION DES EVALUATIONS POUR AMELIORER LA QUALITE DE L'EDUCATION.....	15
2.1. AMPLEUR ET QUALITE DE L'OFFRE D'ENSEIGNEMENT EN AFRIQUE	15
2.2. OBSTACLES A LA REALISATION DES OBJECTIFS DE L'EPT	16
2.3. MOBILISATION RENOUVELEE EN FAVEUR DE L'EPT A DAKAR.....	17
2.4. THEME CENTRAL DU PRESENT DOCUMENT	18
2.5. EXAMENS NATIONAUX (EXTERNES).....	19
2.6. QUALITE DES EXAMENS NATIONAUX EN AFRIQUE.....	21
2.7. AMELIORER LA QUALITE DES EXAMENS NATIONAUX.....	21
2.8. UTILISATION DES DONNEES TIREES DES EXAMENS POUR REFLETER LE NIVEAU SCOLAIRE	22
2.8.1. Rapports de jurys d'examen.....	22
2.8.2. Autres mécanismes de retour d'informations	24
2.8.3. Repérage de catégories d'élèves en difficulté scolaire	25
2.8.4. Limites des données fournies par les examens	25
2.9. UTILISATION DES EXAMENS POUR AMELIORER LE NIVEAU SCOLAIRE DES ELEVES.....	26
2.10. EFFETS DES MODIFICATIONS APPORTEES AUX EXAMENS	26
2.10.1. Une modification des examens peut-elle induire une amélioration du niveau scolaire des élèves ?	27
2.10.2. Une modification des examens peut-elle se traduire par une modification des facultés cognitives des élèves ?	28
2.10.3. Mécanismes destinés à modifier les pratiques scolaires et les résultats des élèves.....	29
2.10.4. Compétition entre écoles	29
2.10.5. Octroi de conseils aux écoles	30
2.10.6. Octroi de soutien à des écoles peu performantes	31
2.10.7. Liens avec les programmes scolaires	32
2.10.8. Problèmes particuliers en cas d'enjeux importants liés aux résultats des évaluations	32
2.10.9. Orientations pour l'amélioration des examens nationaux.....	35
3. EVALUATIONS NATIONALES.....	37
3.1. INFORMATIONS RECHERCHEES DANS LE CADRE DES EVALUATIONS NATIONALES	38
3.1.1. Divers usages des informations fournies par une évaluation nationale.....	39
3.2. DIFFERENCES ENTRE LES EVALUATIONS NATIONALES	39
3.3. INITIATIVES D'EVALUATIONS NATIONALES EN AFRIQUE	40
3.4. IMPACT DES EVALUATIONS NATIONALES	43
3.5. QUESTIONS RELATIVES AUX EVALUATIONS NATIONALES	44
3.5.1. Conditions et pratiques requises pour développer et institutionnaliser les capacités d'exécution d'une évaluation nationale	47

4. EVALUATIONS INTERNATIONALES	49
4.1. EVALUATIONS INTERNATIONALES EN AFRIQUE	49
4.2. DIFFICULTES LIEES AUX EVALUATIONS INTERNATIONALES	50
5. EVALUATIONS A L'ECHELON DE LA CLASSE	52
5.1. QUALITE DE L'EVALUATION A L'ECHELON DE LA CLASSE	53
5.2. AMELIORER LES EVALUATIONS A L'ECHELON DE LA CLASSE.....	54
5.3. UTILISATION DES EVALUATIONS A L'ECHELON DE L'ECOLE DANS DES EXAMENS NATIONAUX	55
6. FACTEURS SUSCEPTIBLES DE FAIRE OBSTACLE A L'UTILISATION DES RESULTATS D'EVALUATIONS ET D'EXAMENS A L'ECHELON DE LA SALLE DE CLASSE	59
6.1. MODE D'INTERVENTION	59
6.1.1. <i>Pertinence des informations</i>	59
6.1.2. <i>Compétences des enseignants</i>	60
6.1.3. <i>Compréhension par l'enseignant des répercussions des changements</i>	60
6.1.4. <i>Complexité de l'enseignement</i>	61
6.1.5. <i>Contexte de la salle de classe</i>	62
6.1.6. <i>Résistance au changement à la perspective des risques qu'il peut présenter pour les élèves.</i>	62
7. CONCLUSION	63
7.1. RECOMMANDATIONS	68
8. REFERENCES BIBLIOGRAPHIQUES.....	70

Encadrés

Encadré 1.	Rapports de jurys d'examen au Swaziland	23
Encadré 2.	Langue d'enseignement	24
Encadré 3.	Inconvénients des palmarès.....	31
Encadré 4.	Incidence des examens sur les enseignants	32
Encadré 5.	Quelques exemples de fraudes à l'examen	34
Encadré 6.	Orientations pour l'amélioration des examens nationaux	35
Encadré 7.	Principaux éléments d'une évaluation nationale	37
Encadré 8.	Différences relatives aux procédures d'évaluations nationales	40
Encadré 9.	Evaluations internationales : le cas de l'Afrique du Sud	51
Encadré 10.	Evaluation à l'échelon de la classe	52
Encadré 11.	Suggestions pour améliorer les évaluations à l'échelon de la salle de classe.....	56
Encadré 12.	Difficultés des enseignants à s'adapter à la réforme d'un examen.....	61
Encadré 13.	Complexité de l'enseignement	61

Acronymes et abréviations

ADEA	Association pour le développement de l'éducation en Afrique
CONFEMEN	Conférence des Ministres de l'Education des pays ayant le Français en partage
MLA	Monitoring Learning Achievement
NESIS	National Education Statistical Information Systems
PASEC	Programme d'Analyse des Systèmes Educatifs des Pays de la CONFEMEN
PRSP	Poverty Reduction Strategy Papers
SACMEQ	Consortium de l'Afrique australe pour le pilotage de la qualité de l'éducation
SAP	Programmes d'ajustements structurels
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture

Remerciements

Nous souhaitons remercier les personnes suivantes pour leur assistance dans la préparation de ce document : R. Agabu, Tunde Adekola, Jean-Marc Bernard, E.D. Bononewe, Jaap Bregman, Vinaygum Chinapah, Paul Coustère, Brigitte Duces, Mourad Ezzine, Birger Fredriksen, Jacques Hallak, Soumana Hamidou, Cathal Higgins, Anil Kanjee, John Khozi, John Lauglo, Lily Mulatu, Paud Murphy, Bob Prouty, Rachidi B. Radji, Ken Ross, Charlotte Sedel, Endale Shewangizaw, Mercy Tembon, Gary Theisen, Adriaan Verspoor, et David Weere. Hilary Walshe a également apporté un appui au niveau du secrétariat.

ABREGE

Le présent document expose les raisons qui justifient l'utilisation des évaluations pour améliorer la qualité de l'éducation et les données d'expérience des systèmes éducatifs africains acquises ces vingt dernières années et ce, pour quatre catégories d'évaluation : (i) les examens nationaux, (ii) les évaluations nationales; (iii) les évaluations internationales et (iv) les évaluations à l'échelon de la salle de classe. Les arguments en faveur d'une réforme des *examens nationaux* reposent en premier lieu sur le constat de la qualité médiocre de ces examens dans maints pays et, en second lieu, sur la conviction que, du fait que les enseignants alignent leur enseignement sur les exigences des examens, l'amélioration des examens induira une modification du contenu enseigné et, par conséquent, de la qualité d'apprentissage des élèves. Plusieurs exemples illustrent la façon dont les résultats d'examen peuvent améliorer les résultats des élèves. Sont également évoqués les effets des réformes des examens sur le contenu de l'enseignement, les résultats scolaires et les facultés cognitives à la lumière d'informations relevées sur le terrain. Diverses initiatives d'évaluation nationale ont été menées sur l'ensemble du territoire africain depuis les années 90. Bien qu'ils s'appliquent à des élèves isolés (choisis comme étant représentatifs de tous les élèves possédant un certain niveau scolaire), les résultats de ces évaluations nationales sont agrégés afin de produire une évaluation du système éducatif. La mission première de ce type d'évaluation est d'apprécier la qualité d'apprentissage des élèves (et de sous-groupes d'élèves identifiables). Pour l'heure, il semble que, dans un certain nombre de pays, les résultats des évaluations ne soient exploités que dans le cadre de commissions gouvernementales, études de synthèse et programmes de réforme de l'éducation. Les *évaluations internationales*, qui présentent de nombreux points communs avec les évaluations nationales au plan des procédures employées, fournissent des données comparatives sur les niveaux scolaires dans différents pays. Quelques pays africains seulement y ont participé. Par contre, certaines évaluations nationales permettent des comparaisons à l'échelon international. Si les *évaluations à l'échelon de la salle de classe* n'ont reçu qu'une attention mineure dans les propositions relatives à l'utilisation des évaluations comme moyen d'améliorer la qualité de l'éducation, il est probable que leur impact sur l'apprentissage des élèves sera plus fort que toute autre forme d'évaluation. Les avantages et les inconvénients d'une prise en compte des évaluations réalisées à l'échelon de l'école dans les notes attribuées aux examens nationaux (système en vigueur dans certains pays) sont examinés. Plusieurs propositions sont faites pour améliorer les examens nationaux, les évaluations nationales et les évaluations à l'échelon de la salle de classe. Enfin, les facteurs susceptibles de faire obstacle à l'utilisation des évaluations pour améliorer la qualité de l'éducation sont étudiés.

1. RESUME

1. Dans le document d'orientation de la Banque mondiale du mois de janvier 1988, intitulé *"Education in Sub-Saharan Africa, Policies for Adjustment, Revitalization, and Expansion"*, l'accent porte sur l'amélioration des niveaux scolaires via la réforme des systèmes d'examen. En juin de la même année, l'association Donors to African Education (aujourd'hui, Association pour le développement de l'éducation en Afrique) a montré son intérêt pour la question des examens avec la création d'un Groupe de travail sur les examens scolaires chargé d'étudier et de promouvoir le rôle des examens dans l'amélioration de la qualité. Par la suite, la Division Education et Formation du Département technique pour l'Afrique de la Banque mondiale a défini les termes de référence pour des études sur les examens existant aux niveaux primaire et secondaire dans les pays d'Afrique subsaharienne. Ces études ont été réalisées avec le soutien financier de la Banque mondiale et du gouvernement irlandais dans six pays anglophones, six pays francophones et deux autres pays. Elles ont donné lieu à la rédaction d'un rapport intitulé *"Using Examinations to Improve Education. A Study in Fourteen African Countries"*, qui a été élaboré par les auteurs du présent document et publié en 1992 par la Banque mondiale.

2. Le présent document dresse un état des lieux des initiatives engagées depuis lors en matière d'évaluation. Il traite des examens nationaux, mais, à la différence du rapport de 1992, il inclut des données nouvelles recueillies (à quelques exceptions près) pour des besoins autres que le présent document. Il réexamine nombre de questions abordées dans le premier rapport en insistant plus spécialement sur le rôle que peuvent jouer des évaluations dans l'amélioration de la qualité d'apprentissage des élèves. Autre différence par rapport au premier rapport : l'analyse sur les effets des évaluations ne se limite pas aux seuls examens nationaux.

3. Outre un regain d'intérêt pour la question des évaluations, la Déclaration mondiale de Jomtien sur l'Education pour tous en 1990 a également mis l'accent sur la nécessité d'instaurer une nouvelle forme d'évaluation – évaluation d'un système ou évaluation nationale - afin de déterminer dans quelle mesure les enfants acquéraient les connaissances théoriques et pratiques, la capacité de raisonnement et les valeurs utiles que les écoles s'engageaient à leur inculquer. Les évaluations nationales constituent le second champ d'évaluation important traité dans ce document. Sont également abordées les évaluations internationales, lesquelles présentent de nombreux points communs avec les évaluations nationales au plan des procédures. Le quatrième type d'évaluation étudié dans ce document concerne les évaluations à l'échelon de la salle de classe.

1.1. Examens nationaux (externes)

4. Les examens nationaux (externes) ont joué un rôle majeur tout au long de l'histoire de l'éducation moderne en Afrique. Ils répondent à plusieurs fonctions dont la plus importante est de sélectionner les élèves jugés aptes à accéder au niveau immédiatement supérieur du système éducatif. En dépit de leur rôle central, la qualité des examens nationaux a été très critiquée. On a en particulier constaté que les domaines de connaissances théoriques et pratiques ainsi évalués étaient limités; qu'ils n'avaient qu'un lointain rapport avec les connaissances théoriques et pratiques dont avaient besoin les élèves dans leur vie quotidienne hors du milieu scolaire; et qu'ils avaient tendance à quantifier le niveau scolaire à un faible échelon taxinomique.

5. Cette situation a des conséquences multiples. En premier lieu, elle pose la question de la validité des examens : ceux-ci tendent à privilégier la vérification des compétences que doivent posséder les élèves désireux de continuer leurs études dans le cycle suivant et ne sont donc pas le reflet des objectifs recherchés par les programmes scolaires, en l'occurrence des acquis des élèves (majoritaires dans la plupart des pays) qui n'entendent pas poursuivre d'études secondaires. En second lieu, dans la mesure où les professeurs concentrent leur enseignement sur les épreuves des examens, lorsque ceux-ci ont des conséquences importantes liées aux performances obtenues, on est en droit de s'interroger sur la nature et la qualité de l'enseignement et de l'apprentissage dans les écoles qui accordent une grande importance aux examens.

6. Ces vingt dernières années, plusieurs pays ont engagé des réformes de leurs systèmes éducatifs afin d'améliorer la qualité des examens : modes d'évaluation plus diversifiés, intégration d'éléments visant à évaluer des aptitudes intellectuelles plus élaborées et évaluation de l'aptitude des élèves à appliquer leurs connaissances théoriques et pratiques dans des contextes extrascolaires et scolaires. L'objectif attendu de ces réformes était d'influer sur le contenu de l'enseignement et sur le contenu de l'apprentissage. Pour renforcer l'impact des examens sur les pratiques appliquées dans les écoles, certains pays y ont ajouté d'autres procédures : publication des résultats d'examen de façon à aiguïser la concurrence entre les établissements scolaires, compte rendu sur les résultats scolaires des élèves, octroi de soutien et de conseils aux écoles peu performantes et création de liens entre les autorités responsables des examens et les autorités responsables de l'élaboration et de la mise en œuvre des programmes scolaires.

7. On ne peut que se réjouir des efforts déployés en faveur d'une amélioration de la qualité des examens et les encourager. Pour autant, la prudence s'impose : il faut éviter de miser sur les seuls examens pour modifier radicalement la qualité d'apprentissage de *tous* les élèves. En premier lieu, s'il est admis qu'une modification du contenu des épreuves induit une modification du contenu enseigné aux élèves dans la salle de classe, les preuves empiriques montrant qu'une modification des examens améliore le niveau scolaire général des élèves ou de leurs facultés cognitives sont peu abondantes. En second lieu, les enjeux élevés qui sont liés aux performances obtenues aux examens peuvent avoir des effets négatifs, s'ils ne sont pas intentionnels, sur les pratiques scolaires indépendamment de la qualité des examens. Le risque est par exemple de réduire le champ des programmes scolaires, de négliger les domaines qui ne figurent pas au programme des épreuves, de privilégier les stratégies d'apprentissage en surface ou à court terme (telles que mémorisation, répétition, apprentissage par cœur), de consacrer beaucoup de temps à la préparation de l'examen et de mettre l'accent sur la motivation extrinsèque de l'apprenant au détriment de sa motivation intrinsèque. Un élément est particulièrement significatif dans le contexte de l'Education pour tous : du fait que leur réputation dépend des résultats obtenus par leurs élèves aux examens, les enseignants ont tendance à s'intéresser en priorité aux élèves qui offrent le meilleur potentiel de réussite. D'où le risque d'une augmentation des taux de redoublement et d'abandon scolaire.

1.2. Evaluations nationales

8. Contrairement aux examens nationaux qui sont depuis longtemps une caractéristique inhérente aux systèmes éducatifs, les évaluations nationales (parfois appelées évaluations du système, évaluations d'apprentissage et, de façon moins correcte, évaluations des résultats d'apprentissage) sont relativement nouvelles. Par évaluation nationale, on entend un procédé utilisé pour illustrer le niveau, non pas d'élèves individuellement, mais de l'ensemble d'un système éducatif ou d'une partie clairement définie d'un système (les élèves en quatrième année de primaire ou les élèves âgés de 11

ans par exemple). La raison pour laquelle les évaluations nationales ont été instaurées est que les données de l'éducation sur les intrants éducatifs que l'on collectait par le passé étaient bien souvent peu pertinentes ou difficilement utilisables par les planificateurs de l'éducation. Les évaluations nationales avaient pour but de résoudre ce problème en apportant des informations sur les « produits » ou les « résultats » de la scolarité (niveau des élèves, inégalités du système par exemple), informations susceptibles d'être exploitées en lien avec les données de départ afin de servir de base au processus d'élaboration de politiques et de prise de décision dans le secteur de l'éducation.

9. Les principales questions traitées dans le cadre des évaluations nationales sont les suivantes : Quelle est la qualité d'apprentissage des élèves ? Existe-t-il des points forts ou des points faibles dans leurs connaissances théoriques et pratiques ? Existe-t-il des écarts entre les résultats de différents sous-groupes de la population ? Y a-t-il un lien entre les résultats scolaires et les caractéristiques du contexte d'apprentissage ? Les performances des élèves varient-elles avec le temps ?

10. Durant les années 90 en Afrique, de nombreuses initiatives sont nées dans le domaine des évaluations nationales. On en compte quatre grandes catégories. Trois d'entre elles concernent des initiatives similaires menées dans divers pays : le projet "Monitoring Learning Achievement" (MLA); le projet du Consortium de l'Afrique australe pour le pilotage de la qualité de l'éducation (SACMEQ); le Programme d'Analyse des Systèmes Educatifs des Pays de la CONFEMEN (PASEC). La quatrième regroupe des évaluations nationales indépendantes réalisées dans quelques pays. A ce jour, des évaluations MLA ont été pratiquées dans 47 pays africains, des évaluations SACMEQ dans 15 pays et des évaluations PASEC dans neuf pays.

11. Ces évaluations présentent des points communs. Ce sont des évaluations portant sur les politiques, auxquelles sont associés des planificateurs et des gestionnaires et qui ont pour but de fournir des informations utiles à l'élaboration de politiques. Elles évaluent le niveau des élèves dans des disciplines de base. Elles utilisent la technique d'échantillonnage. Le renforcement des capacités et du couple décideurs/chercheurs compte parmi leurs objectifs prioritaires. Elles s'appuient sur un modèle du système éducatif de type entrées-sorties et visent à dégager les paramètres qui interviennent dans les résultats scolaires.

12. Pour tous les programmes d'évaluation, un compte rendu des résultats a été diffusé dans plusieurs secteurs : débats et formulation de politiques, études de synthèse sur la politique éducative, études du secteur éducatif à l'échelon national, l'objectif étant de réexaminer l'adéquation des ressources et d'apporter une aide décisionnelle aux responsables des orientations politiques. Plusieurs pays ont procédé à trois évaluations nationales indépendantes (respectivement parrainées ou financées par un organisme différent). Cela laisse supposer que les évaluations n'étaient pas une réponse à des besoins perçus à l'échelon local et qu'elles ne font pas partie intégrante des structures et activités normales des ministères. Compte tenu de la nécessité d'une telle intégration, du coût de ces opérations et des problèmes relatifs au recrutement de personnel compétent pour les mener à bien, l'urgence d'une rationalisation se fait clairement sentir.

1.3. Evaluations internationales

13. Les évaluations internationales diffèrent des évaluations nationales dans la mesure où elles entendent situer le niveau scolaire des élèves d'un pays par rapport à celui des élèves d'un autre pays. Elles donnent en outre des indications sur les différentes

manières d'aborder les mêmes disciplines dans les programmes scolaires d'un pays à l'autre.

14. Un petit nombre seulement de pays africains a participé à des évaluations internationales. Néanmoins, si les évaluations MLA, PASEC et SACMEQ ont par essence une portée nationale, elles permettent d'effectuer des comparaisons internationales sur la base des comptes rendus ainsi publiés.

1.4. Evaluations à l'échelon de la salle de classe

15. Evaluer l'apprentissage des élèves à l'échelon de la salle de classe (évaluation à la fois par les enseignants et par les élèves eux-mêmes) fait partie intégrante du processus d'enseignement-apprentissage. Cette forme d'évaluation revêt, pour une large part, un caractère subjectif, informel, immédiat, continu et intuitif, dans la mesure où elle interagit avec l'apprentissage au fur et à mesure de son avancement, le pilotage du comportement de l'élève, les performances scolaires et la réactivité à l'enseignement. Son but est de déterminer le niveau de connaissances théoriques, de connaissances pratiques ou de compréhension des élèves à un instant donné, de diagnostiquer les difficultés qu'ils sont susceptibles de rencontrer, de prendre des décisions quant aux prochaines étapes pédagogiques à accomplir (revoir certains points du programme ou poursuivre) et d'évaluer l'apprentissage qui s'est fait pendant un cours.

16. Il est manifeste que la qualité des pratiques d'évaluation utilisées par les enseignants laisse parfois à désirer à de nombreux égards. Parmi les problèmes constatés, on peut citer la mauvaise formulation des questions, la prédominance de questions qui appellent des réponses courtes portant sur des connaissances factuelles, l'évocation de réponses qui font plus appel à la répétition qu'à la réflexion et l'absence de procédures conçues pour développer les facultés cognitives supérieures des élèves.

17. Hormis quelques exceptions notables dans les pays francophones, l'attention portée dans les réformes à l'amélioration des évaluations à l'échelon de la salle de classe a été réduite. Pourtant, étant donné leur rôle central dans le processus d'enseignement-apprentissage, améliorer les pratiques d'évaluation à l'échelon de la salle de classe mériterait de constituer l'une des priorités de toute réforme visant à favoriser l'apprentissage des élèves.

18. Plusieurs pays africains ont introduit, ou prévoient de le faire, un élément d'évaluation à l'échelon de l'école dans leurs examens nationaux. En effet, notamment dans l'enseignement primaire, d'aucuns espèrent à terme remplacer complètement les examens externes par des évaluations à l'échelon de l'école, considérant qu'elles sont la seule formule permettant d'évaluer valablement l'ensemble des compétences spécifiées dans les programmes scolaires et d'éviter les effets négatifs des examens externes sur l'enseignement et l'apprentissage. Au plan concret, l'intégration d'une évaluation locale dans des examens nationaux s'est toutefois révélée problématique. En dépit des efforts investis et d'une forte motivation en faveur de cette formule, les difficultés pratiques ont plus d'une fois conduit à abandonner le projet, à en différer l'exécution ou à réduire la part de cette évaluation à l'échelon de l'école à la portion congrue, sinon au rang de symbole.

1.4.1. Problèmes liés à l'utilisation des informations tirées des évaluations à l'échelon de la salle de classe

19. Transmettre des informations aux enseignants et induire des changements dans leur comportement est une tâche bien difficile. L'espoir que les informations tirées

des évaluations et des examens (via la publication des résultats, les bulletins d'informations, les ateliers) aboutissent à un changement radical de la culture scolaire et à une augmentation sensible du niveau de tous les élèves doit être tempéré par la prise en compte de divers facteurs susceptibles de rendre vaines les intentions des réformateurs. À défaut d'identifier ces facteurs et de les traiter, les politiques menées en matière d'évaluation risquent d'avoir un impact très limité sur la qualité d'apprentissage des élèves. Les facteurs analysés sont les suivants : mode d'intervention, pertinence des informations, compétences de l'enseignant, compréhension par l'enseignant des répercussions des changements, complexité de l'enseignement, contexte de la salle de classe et résistance des enseignants et des parents au changement compte tenu des risques qu'il peut présenter pour les élèves.

1.5. Mécanismes d'aide à l'utilisation des évaluations

20. L'une des conclusions de l'analyse présentée dans ce document est que, si les résultats des évaluations *peuvent* améliorer la politique et la gestion des ressources éducatives et peuvent façonner les pratiques pédagogiques des enseignants, le succès n'est pas assuré. Ce succès dépendra tout d'abord de la volonté politique des gouvernements de soutenir ces efforts. Ces efforts devront s'exercer à deux niveaux. En premier lieu, il conviendra d'institutionnaliser les systèmes d'évaluation nationale et de les intégrer dans les structures et processus d'élaboration des politiques et de prise de décision des dirigeants. Il importera en outre de veiller à aligner les évaluations nationales sur d'autres grands principes d'orientation pédagogique du système éducatif – autres systèmes d'évaluation (y compris l'alignement des normes de niveau), programmes scolaires, formation des maîtres, renforcement des capacités des écoles et mesures visant à supprimer les injustices. Des actions devront en outre être engagées à l'échelon de l'école. Il conviendra d'ouvrir des circuits de communication adaptés pour informer les enseignants des effets à attendre des conclusions des évaluations. La tâche risque d'être difficile, notamment s'agissant de problèmes touchant à des objectifs pédagogiques, à l'équité et à la qualité; dans ce cas, les enseignants pourront avoir besoin d'un soutien considérable et durable pour interpréter les réformes et concevoir des stratégies d'enseignement adaptées.

2. CONTEXTE DES PROPOSITIONS RELATIVES A L'UTILISATION DES EVALUATIONS POUR AMELIORER LA QUALITE DE L'EDUCATION

21. La Conférence mondiale sur l'Education pour tous (EPT) qui s'est tenue à Jomtien en Thaïlande en 1990 laissait espérer une amélioration de l'équité et de la qualité des systèmes éducatifs dans le monde dès lors que les pays chercheraient à garantir leur compétitivité économique dans une économie mondiale en marche vers le 21^{ème} siècle. Non seulement l'accès aux offres de formation devrait être ouvert à tous les enfants, mais ceux qui en bénéficient devraient acquérir des connaissances théoriques et pratiques, une capacité de raisonnement et des valeurs utiles (UNESCO, 1990). Depuis lors, le niveau scolaire des élèves est devenu un critère clé pour juger de la qualité de l'éducation. Les recommandations élaborées lors de la conférence ont eu un retentissement considérable, car elles étaient l'expression d'une convergence de vues entre des gouvernements, des organisations et des bailleurs de fonds. Ainsi, de nombreux pays dans le monde s'en sont fait l'écho concrètement en accordant la priorité à l'éducation de base, en favorisant l'égalité d'accès des filles et des garçons à l'éducation, en améliorant l'accès des populations rurales les plus pauvres à l'éducation, en améliorant la qualité de l'éducation à tous les échelons, en s'attaquant au problème de la pertinence pratique des programmes scolaires par rapport à la vie des élèves et en mobilisant leurs efforts en faveur d'un renforcement des capacités.

22. A bien des égards, les aspirations de la conférence n'avaient aucun caractère de nouveauté particulier. Entre les années 60 et 80, nombre de pays en développement ont proposé des programmes d'amélioration de l'enseignement primaire axés vers le développement quantitatif et l'amélioration qualitative. Ces programmes prévoyaient notamment d'instaurer un meilleur équilibre entre les disciplines théoriques et les disciplines pratiques/techniques, d'inclure des activités scolaires et communautaires, d'améliorer la préparation des enseignants, d'étendre les services d'inspection, de contrôle et de conseil pour les enseignants, d'améliorer les méthodes d'évaluation des élèves, d'élargir la mise à disposition des manuels scolaires et d'accroître l'efficacité de la planification de l'éducation (Grisay & Mähllck, 1991). Cependant, l'originalité de la déclaration de Jomtien par rapport à ces programmes résidait, non pas seulement dans sa dimension internationale et inter institutionnelle, mais aussi dans le fait qu'elle obligeait à définir l'amélioration de la qualité ou à en fixer les objectifs et qu'elle soulignait la nécessité d'évaluer les effets des réformes sur le niveau scolaire.

2.1. Ampleur et qualité de l'offre d'enseignement en Afrique

23. Il est facile de dire, rétrospectivement, que l'objectif de la Déclaration sur l'Education pour tous, à savoir garantir une éducation universelle de base dans tous les pays d'ici à l'horizon 2000, était démesurément ambitieux. Cet objectif n'a pas été atteint. Certains pays ont pris de l'avance, d'autres ont préservé leurs acquis sans opérer de percée décisive, d'autres encore ont régressé (S. Nalovu cité dans ADEA, 1999). Si le taux net de scolarisation a progressé en Afrique subsaharienne, la population d'enfants d'âge scolaire non scolarisés a pour sa part augmenté depuis 1980, après avoir diminué dans les années 70 (UNESCO, 2000a). L'ampleur de la tâche qui reste à accomplir est clairement perceptible quand on considère le fait qu'en Afrique, 42,4 millions d'enfants d'âge scolaire

ne fréquentent pas l'école (UNESCO, 2002). Toutefois, scolariser tous les enfants n'est qu'une première étape vers la réalisation de l'objectif d'une Education pour tous : en effet, l'axe majeur de toute réforme doit être l'acquisition et les résultats de l'apprentissage, et non pas exclusivement la scolarisation. Là encore, l'ampleur de la tâche apparaît clairement quand on analyse la qualité de l'enseignement dispensé à ceux qui fréquentent l'école. D'une conférence régionale sur l'Education pour tous en Afrique subsaharienne qui s'est tenue à Johannesburg en décembre 1999 et à laquelle ont participé des ministres de l'éducation et des représentants de la société civile et d'organisations internationales de développement, il est ressorti que : pour ceux qui étaient intégrés dans le système éducatif, le contenu de l'enseignement est de qualité « médiocre et les programmes scolaires souvent déconnectés des besoins des apprenants et des exigences du progrès social, culturel et économique » (UNESCO, 2000b, p. 26); « la planification et la capacité de gestion de l'éducation... sont dans une large mesure insuffisamment développées » (p. 26); « seule une faible proportion d'enfants atteint le niveau minimum requis de compétences et nos systèmes éducatifs ne sont pas à la hauteur des attentes » (p. 28).

24. S'agissant de l'éducation dans divers pays du continent, les observations n'apportent aucune note d'optimisme dans ce sombre tableau. A titre d'exemple, on peut lire que le système éducatif au Ghana est d'une qualité déplorable, comme le démontre les « résultats catastrophiques des tests nationaux par critères » réalisés entre 1992 et 1996 (N'tchougan-Sonou, 2001). En Ethiopie en 1990, plus de 20 % des candidats à l'examen de fin de 6^{ème} année, 40 % à l'examen de fin de 8^{ème} année et 70 % à l'examen de fin de cycle secondaire ont échoué. Et ce, en dépit du fait qu'environ 20 % des élèves de 1^{ère} et de 9^{ème} année étaient des redoublants. Au Mozambique, un tiers des élèves de 5^{ème} année et la moitié des élèves de 7^{ème} année ont échoué aux examens nationaux (Takala, 1998). Au Sénégal, entre 1990 et 1995, les taux de réussite à l'examen de fin de cycle primaire ont été inférieurs à 30 % en moyenne (ADEA, 2002). En Guinée, seuls 6 % des élèves ont obtenu leur « maîtrise » de français et aucun en mathématiques (5 % ont obtenu leur « maîtrise partielle ») lors des épreuves destinées à être le reflet des objectifs recherchés par les programmes scolaires (Carron & Châu, 1996).

2.2. Obstacles à la réalisation des objectifs de l'EPT

25. Malheureusement, les efforts investis dans l'application de politiques axées vers l'Education pour tous sont intervenus à une époque où de nombreux pays du continent étaient en proie à la guerre civile, aux catastrophes naturelles et à la pandémie de VIH/SIDA. Qui plus est, ils ont été lancés au moment où étaient mis en œuvre des Programmes d'ajustements structurels (SAP), d'où la vulnérabilité des dépenses d'éducation. Dans le cas de l'éducation, les mesures des SAP visaient pour la plupart à améliorer l'efficacité (réduction des coûts unitaires) et la rentabilité de l'enseignement secondaire (via une augmentation des frais de scolarité et l'encouragement de l'offre privée), autrement dit des mesures qui ne sont guère propices à une amélioration de la qualité parallèlement à un élargissement de l'offre. Toutefois, « à terme, que l'éducation contribue ou non au développement dépend moins du nombre d'enfants qui achèvent leur scolarité que de ce qu'ils y apprennent ou n'apprennent pas » (Samoff, 1999, p. 199).

26. Face à la pénurie de ressources humaines et financières disponibles pour mener à bien l'objectif de l'EPT, on pourrait s'attendre à ce que l'essor considérable de l'éducation en Afrique durant les dernières décennies se soit accompagné d'une dégradation de sa qualité (voir Bude, 1993). Comment escompter par exemple de l'Ouganda qu'il parvienne à maintenir, encore moins accroître, la qualité de son système éducatif alors que les effectifs scolaires en primaire progressent de 2,5 millions (en 1996) à 6,5 millions (en 1999) et le taux d'encadrement de 38 pour 1 à 62 pour 1 (ADEA, 1999) ? Le cas s'est produit en Zambie. Constitué au lendemain de Jomtien, le Groupe de

travail national *Focus on Learning* a constaté que le « prix à payer pour un développement quantitatif rapide au cours des premières années a été une sérieuse dégradation de la qualité de l'éducation scolaire » (cité dans Takala, 1998). Dans plusieurs pays toutefois, il semble que, comme dans le projet initial relatif à l'Education pour tous, on n'ait guère prêté attention aux éventuels compromis à prévoir dans le cadre des politiques éducatives entre développement quantitatif et amélioration qualitative. Face à la chute spectaculaire de la qualité de l'enseignement, au recrutement d'enseignants moins qualifiés, à l'augmentation des effectifs des classes, au manque de moyens et à la pénurie grandissante de manuels scolaires, de nombreux parents en sont donc venus à douter de la valeur de l'éducation dispensée à leurs enfants (Chapman & Mählck, 1993).

2.3. Mobilisation renouvelée en faveur de l'EPT à Dakar

27. Les progrès accomplis durant les années 90 sont, il faut le reconnaître, décevants. En l'an 2000, le Cadre d'action de Dakar a été l'occasion pour les pays, les organisations et les bailleurs de fonds de réaffirmer leur détermination à œuvrer en direction d'une nouvelle série d'objectifs relatifs à l'Education pour tous (UNESCO, 2000a). Ces objectifs étaient notamment :

- de garantir que, d'ici à 2015, tous les enfants, en particulier les filles, les enfants vivant dans des conditions difficiles et les enfants appartenant à des minorités ethniques, aient accès à un cursus complet d'éducation primaire de bonne qualité (objectif 3);
- d'éliminer les inégalités de scolarisation dans l'enseignement primaire et secondaire entre les garçons et les filles d'ici à 2005 et de parvenir à l'égalité des sexes dans l'éducation d'ici à 2015, tout en mettant l'accent sur la nécessité d'assurer l'égalité d'accès et de chances de toutes les filles à une éducation de base de qualité (objectif 5); et
- d'améliorer tous les aspects de la qualité de l'éducation dans un souci d'excellence pour tous de façon à obtenir pour tous des résultats d'apprentissage reconnus et quantifiables, notamment en ce qui concerne la lecture, l'écriture, le calcul, ainsi que les compétences indispensables dans la vie courante (objectif 6) (UNESCO, 2002, p. 13).

28. Contrairement à la déclaration de Jomtien, le Cadre de Dakar a défini un plan d'actions destiné à conférer un sentiment d'urgence et à créer un climat de transparence (UNESCO, 2002). A l'évidence, les politiques et les pratiques adoptées dans le droit fil des objectifs de l'EPT dépendront du contexte local et varieront d'un pays à l'autre. Cependant, comme tous les pays se sont engagés à garantir l'obtention, pour tous, de résultats d'apprentissage reconnus et quantifiables, une certaine forme d'évaluation devra être mise en place afin de déterminer si la situation s'est, ou non, améliorée.

29. La nécessité d'évaluer les résultats des élèves implique de ne plus évaluer la qualité de l'éducation en se référant principalement aux intrants (taux de participation des élèves, infrastructures, programmes scolaires, livres et formation des maîtres par exemple), sauf dans des cas bien précis (Grisay & Mählck, 1991). Certes, la pertinence des intrants dans l'évaluation de la qualité ne fait pas de doute. Pour autant, affirmer que les intrants d'un haut niveau de qualité sont invariablement liés à de bons résultats scolaires n'est pas toujours vrai. Et à l'inverse, dire que les intrants d'un faible niveau de qualité sont toujours liés à de médiocres résultats scolaires ne l'est pas plus.

30. Ce regain d'intérêt pour l'évaluation de l'apprentissage des élèves au cours de leur scolarité ne s'explique pas exclusivement par le besoin de disposer d'informations adaptées pour analyser l'effet des politiques d'EPT. Il ne se limite pas non plus aux seuls pays en développement. Ces dernières années, les pays industrialisés et les pays en développement ont manifesté un net intérêt pour l'élaboration d'une approche corporative de l'administration, de pair avec une progression du « managérialisme ». Cette approche est inspirée de concepts provenant du monde des affaires, telles que la planification stratégique et opérationnelle, le perfectionnement constant, l'importance accordée aux « livraisons demandées » et aux résultats, l'assurance de la qualité, les systèmes d'incitation et de responsabilisation par rapport aux résultats et le concept du citoyen consommateur. Dans cette perspective, les performances sont définies en termes de résultats, des objectifs de performances sont fixés, le degré de réalisation des résultats est établi sur la base d'indicateurs de performances et les décisions concernant les allocations de ressources s'appuient sur des données relatives aux performances. La finalité des activités est de permettre un retour d'informations rapide et continu concernant un nombre limité de résultats jugés importants pour des responsables politiques, des décideurs, des administrateurs et des partenaires (Kellaghan, 2003; Kellaghan & Madaus, 2000).

2.4. Thème central du présent document

31. Le thème central du présent document s'articule autour de l'utilisation des évaluations pour le pilotage et l'évaluation de l'apprentissage. Reconnaître la nécessité de piloter ce que les élèves apprennent conduit à quatre autres nécessités plus spécifiques : (i) la nécessité d'obtenir un état descriptif des connaissances théoriques et pratiques constitutives d'une « éducation de qualité » ; (ii) la nécessité de disposer de moyens techniques et de capacités d'organisation permettant de quantifier les résultats des élèves; (iii) la nécessité de disposer de moyens techniques permettant d'évaluer les progrès accomplis vers l'objectif d'une « éducation de qualité » ; et (iv) la nécessité de posséder un savoir-faire permettant de traduire les résultats des évaluations en politiques et en méthodes pédagogiques propres à améliorer la qualité de l'apprentissage (UNESCO, 2000b).

32. Le terme « évaluation » sera employé pour désigner toute procédure ou activité visant à collecter des données sur les connaissances théoriques, les connaissances pratiques ou les attitudes d'un apprenant ou d'un groupe d'apprenants (Kellaghan & Greaney, 2001). Quatre grandes catégories d'évaluation seront analysées : examens nationaux (externes) ; évaluations nationales ; évaluations internationales ; et évaluations à l'échelon de la salle de classe. Il est à noter qu'à chaque catégorie correspond une grande diversité de pratiques. Cette classification est cependant justifiée par le fait que les procédures administratives varient d'une catégorie à l'autre et que les finalités fondamentales des évaluations diffèrent selon les catégories.

33. Conscients des difficultés passées des gouvernements en la matière faute de disposer de données suffisantes pour permettre la prise de décisions, notre but sera de définir pour chaque catégorie dans quelle mesure les informations tirées des évaluations peuvent aider les décideurs à améliorer la gestion et, partant, la qualité du système éducatif (voir Chapman & Mählck, 1993). Comme on le verra par la suite, les résultats des évaluations peuvent être utilisés à d'autres fins que l'aide décisionnelle en matière de gestion. Lorsque les évaluations ont des conséquences importantes liées aux performances (par exemple si les résultats obtenus à un examen public sont déterminants pour l'obtention d'un diplôme, le passage dans l'année supérieure ou la sélection en vue d'accéder à des études ultérieures ou à un emploi), on peut s'attendre à ce qu'elles aient une incidence directe sur l'enseignement et sur l'apprentissage ; à ce titre, on se doit par conséquent de les considérer comme un moyen d'améliorer les résultats des élèves.

34. Le principal intérêt des évaluations dans le contexte de l'Education pour tous a été, non pas tant d'évaluer les résultats des élèves que de contrôler *les performances du système éducatif*. Cette question sera abordée dans les chapitres consacrés aux évaluations nationales et internationales. Toutefois, compte tenu du rôle important que l'évaluation *d'élèves isolés* dans des examens nationaux (externes) joue dans les systèmes éducatifs en Afrique, on examinera également la manière dont les examens pourraient servir à façonner les pratiques pédagogiques des enseignants selon certains modèles et, par là même, à améliorer la qualité d'apprentissage des élèves.

35. Les évaluations à l'échelon de la salle de classe n'ont reçu qu'une attention mineure dans les propositions relatives à l'utilisation des évaluations pour améliorer l'apprentissage des élèves. L'une des raisons probables est qu'il est difficile pour les décideurs d'influer sur les pratiques des enseignants dans leur classe sur une courte période ou sans qu'il soit nécessaire de leur dispenser une formation continue nécessairement onéreuse. Il est plus simple et moins coûteux de recourir à un système d'évaluation centralisé existant (examens nationaux) ou de commanditer une évaluation externe (évaluation nationale). Toutefois, comme l'évaluation de l'apprentissage des élèves en salle de classe par des enseignants occupe une place centrale dans le processus d'apprentissage de tous les élèves, elle n'en mérite pas moins qu'on y prête attention lorsque l'on étudie des propositions visant à utiliser les évaluations comme moyen d'améliorer cet apprentissage.

2.5. Examens nationaux (externes)

36. Les examens nationaux (externes) ont joué un rôle majeur tout au long de l'histoire de l'éducation moderne en Afrique. La plupart des pays fonctionnent selon le principe de trois grands examens gérés par un organisme extérieur à l'école (en général, autorité nationale chargée des examens dans les pays anglophones et ministère de l'éducation dans les pays francophones). Le premier examen se déroule à la fin de la scolarité primaire et a pour but de vérifier les connaissances acquises par les élèves dans les principales matières figurant au programme (anglais ou français, langue locale, matières scientifiques et sciences humaines, par exemple). Dans la plupart des cas, les notes obtenues à cet examen conditionnent l'entrée de l'élève en cycle secondaire. Après deux ou trois ans de scolarité secondaire, les élèves passent un autre examen (qui porte ordinairement sur un éventail plus large de matières) dont les résultats servent à déterminer si les élèves sont ou non capables d'accéder au second cycle d'enseignement secondaire. Au terme du cycle secondaire, un troisième examen public est prévu; les résultats sont déterminants pour la poursuite d'études supérieures et pour le projet professionnel des élèves.

37. En Afrique, les examens répondent à plusieurs fonctions importantes qui reflètent le contexte social et pédagogique dans lequel ils se déroulent. En premier lieu, ils sont un moyen de contrôler les éléments hors norme du système éducatif et de veiller à ce que toutes les écoles appliquent les mêmes règles en matière d'enseignement; cette pratique était indispensable à l'époque coloniale où les écoles étaient en majorité régies selon un mode de gestion privée. En second lieu, ils servent à sélectionner les élèves dans des systèmes éducatifs pyramidaux où le nombre de places diminue à mesure que l'on franchit des niveaux de plus en plus élevés. En troisième lieu, les examens ont une fonction d'authentification au travers de la délivrance de diplômes/certificats, même si cet aspect est souvent occulté par leur rôle de sélection. L'authentification officielle des résultats scolaires peut, cependant, être importante pour des élèves qui souhaitent entrer sur le marché du travail ou suivre une formation; mais les certificats de bas niveau perdent leur valeur sur le marché du travail du fait que le nombre de leurs titulaires augmente. Enfin, les examens nationaux peuvent avoir une fonction de transparence au

niveau des enseignants et des écoles. C'est le cas en particulier lorsque les résultats obtenus par les élèves aux examens sont rendus publics (Kellaghan, 1992; Kellaghan, & Greaney, 1992; Omolewa & Kellaghan, 2003).

38. Plusieurs facteurs expliquent la persistance et le rôle clé des examens nationaux en Afrique. Ils sont perçus comme un moyen de distribuer des avantages éducatifs peu abondants de façon objective et impartiale; il est néanmoins vrai que le risque de discrimination à l'encontre de minorités, de populations rurales, des filles et d'élèves dont la langue maternelle n'est pas celle utilisée pour l'examen existe. Ils servent à spécifier des normes et des objectifs précis pour les enseignants et les élèves. Ils peuvent être utilisés pour étayer des changements de programmes et de méthodes pédagogiques et pour préserver les normes nationales. Enfin, notamment en fin de cycle secondaire, ils légitiment l'appartenance à une société mondiale internationale et facilitent la mobilité internationale.

39. Dans ce chapitre, nous examinerons tout d'abord la qualité des examens nationaux en Afrique. En dépit des différences qui existent d'un système à l'autre et des efforts déployés dans maints pays au cours des vingt dernières années pour améliorer la qualité, ces examens présentent de nombreuses lacunes. Ces lacunes sont, pour partie, inhérentes à tout examen final écrit, c'est-à-dire de type papier-crayon. Dans d'autres cas, des améliorations sont envisageables tout en conservant le modèle d'épreuve papier-crayon. Après un rapide exposé sur la qualité des examens, nous aborderons la question de l'utilisation des données tirées des examens comme illustration du niveau scolaire. L'analyse des résultats des élèves peut fournir des indications sur la nature des performances des élèves à la fin d'un cursus (par exemple matières du programme dans lesquelles ils obtiennent de mauvais résultats) et sur le mode de répartition de ces résultats entre garçons et filles, d'une école à l'autre ou d'une région à l'autre. Nous présenterons ensuite la raison d'être des propositions visant à utiliser les examens comme levier de réformes, ainsi que les quelques preuves empiriques qui montrent l'impact d'une réforme des examens sur le contenu de l'enseignement dispensé dans les salles de classe et sur les résultats des élèves. Nous exposerons également plusieurs mécanismes visant à utiliser les données d'évaluation pour instaurer un changement de pratiques et relever les niveaux scolaires. Après avoir évoqué les problèmes qui peuvent se poser lorsque des examens comportant un enjeu élevé sont utilisés pour influencer sur l'enseignement et l'apprentissage, nous proposerons un certain nombre d'orientations destinées à améliorer la qualité des examens.

40. Il convient de noter que notre commentaire porte sur le système éducatif formel. Cela ne signifie nullement que le secteur non formel ne constitue pas un élément important de l'effort éducatif en Afrique. On trouve en effet des élèves de ce secteur qui se préparent aux examens nationaux. Dans de nombreux systèmes éducatifs, il y a de beaucoup de candidats « externes » qui ne suivent pas les cours des écoles reconnues. Toutefois, il s'agit surtout d'élèves qui ont moins bien réussi que ce qu'ils espéraient obtenir lors de tentatives précédentes à un examen et qui ne suivaient pas des cours qui seraient considérés comme relevant du secteur informel. Que les examens nationaux ne soient pas perçus comme pertinents pour l'éducation non formelle est en grande partie dû aux différences qui existent entre les objectifs de l'éducation formelle et de l'éducation non formelle, aux différences de personnalité des élèves de chaque secteur et au fait que beaucoup d'enseignants du secteur non formel ne souhaitent pas s'impliquer dans des évaluations formelles des résultats scolaires. Bien que la plupart des élèves du secteur non formel ne souhaitent pas entrer dans le système formel, pour ceux qui le souhaitent, il conviendrait de prévoir des dispositions formelles afin d'établir des équivalences entre les qualifications obtenues dans le cadre formel et le cadre non formel.

2.6. Qualité des examens nationaux en Afrique

41. Nombre d'études et de rapports officiels font état des déficiences des examens actuels. Tout d'abord, les examens en primaire comme en secondaire se limitant pour la plupart à des épreuves écrites de type papier- crayon, ils ne tiennent pas compte de certains domaines de connaissances théoriques et pratiques qui, bien qu'étant au programme, ne sont pas quantifiables dans le cadre d'épreuves de ce type. Deuxième point : les examens mettent l'accent sur l'acquisition du savoir cognitif (en particulier langue et mathématiques en fin de cycle primaire) et n'accordent que très peu d'attention au savoir pratique. Troisième point : il semble que, pour nombre d'examens, les questions posées tendent à mesurer les résultats à un faible niveau taxinomique (appel à la mémorisation ou à la reconnaissance de connaissances factuelles) plutôt que les compétences dont les élèves ont besoin dans leur vie future, comme la capacité à faire la synthèse de documents ou l'application de connaissances à de nouvelles situations (tirer des conclusions, élaborer une séquence logique d'étapes dans la résolution d'un problème, bâtir une argumentation). Quatrième point : de nombreux examens n'ont qu'un rapport lointain avec la vie quotidienne des élèves en dehors de l'école. Ils portent principalement sur des sujets et des applications à caractère scolaire et ne cherchent pas, par exemple, à déterminer si un élève sait gérer son budget ou résoudre des problèmes de santé chez lui. Cinquième point : la qualité des questions est souvent médiocre : elles sont mal formulées, les choix proposés dans les questionnaires à choix multiple laissent à désirer ou les critères de notation manquent de précision (ERGESE, 1986 ; Kellaghan & Greaney, 1992 ; Kelly, 1991 ; Little, 1982 ; Oxenham, 1983 ; Somerset, 1996).

42. Cette situation a deux conséquences majeures. En premier lieu, elle pose la question de la validité des examens : ceux-ci tendent à privilégier la vérification des compétences que doivent posséder les élèves désireux de continuer leurs études dans le cycle suivant et ne sont donc pas le reflet des objectifs recherchés par les programmes scolaires, en l'occurrence, des acquis des élèves (majoritaires dans la plupart des pays) qui n'entendent pas poursuivre d'études secondaires. En second lieu, dans la mesure où les professeurs concentrent leur enseignement sur les épreuves de l'examen lorsque ceux-ci ont des conséquences importantes liées aux performances obtenues, on est en droit de s'interroger sur la nature et la qualité de l'enseignement et de l'apprentissage dans les écoles. Nous reviendrons sur cette question ultérieurement.

2.7. Améliorer la qualité des examens nationaux

43. Au fil des années, diverses tentatives ont été menées dans maints pays pour améliorer la qualité des examens. Il s'est agi en général d'améliorations à caractère technique, axées vers une plus grande efficacité. Face à l'augmentation du nombre d'élèves et de copies d'examen à corriger, le système de questionnaires à choix multiple a été adopté par de nombreux pays. Certains utilisent une formule mixte, questionnaires à choix multiple et réponses courtes; d'autres ont adopté les questionnaires à choix multiple, puis ont abandonné la formule pour la remplacer par le système de réponses courtes (Bude & Lewin, 1997; Lewin & Dunne, 2000).

44. D'autres réformes impliquent une modification beaucoup plus radicale des examens, notamment en termes de contenu. Au cours des années 80 et 90, de nombreux pays ont entrepris de réformer leur système d'examens, parmi lesquels l'Ouganda, le Kenya (1985), le Zimbabwe (1990), la Tanzanie (1994) (Bude, 1997) et la Namibie (1991). A la suite de la publication du document d'orientation de la Banque mondiale (1988) intitulé "*Education in Sub-Saharan Africa*", qui préconisait « une mobilisation renouvelée à l'égard des niveaux scolaires, essentiellement par le biais d'un renforcement des systèmes d'examen » (p. 93), la première réunion plénière de l'association Donors to

African Education (DAE) (aujourd'hui Association pour le développement de l'éducation en Afrique – ADEA), reflétant l'état d'esprit de nombreux pays et les tendances nouvellement apparues en la matière, a évoqué la nécessité d'améliorer la qualité de l'éducation. Un Groupe de travail chargé des examens scolaires (WGSE, "Working Group on School Examinations") a alors été constitué, attestant par là du rôle clé que pourraient jouer les examens dans l'amélioration de la qualité; puis, en 1988 et 1989, quatorze pays ont entrepris des études sur les examens (Kellaghan & Greaney, 1992). Les objectifs de ces études étaient de contribuer à améliorer la qualité de l'éducation par des moyens rationnels et efficaces et d'aider les pays subsahariens à renforcer leurs capacités institutionnelles en ce sens.

45. Parmi les réformes des examens qui ont été préconisées ici ou là, on peut citer : élargir le champ des examens afin de refléter les programmes scolaires ; employer différents modes d'évaluation (écrite, pratique, orale) ; réaménager les examens afin d'y inclure des éléments permettant de mesurer les aptitudes intellectuelles plus élaborées; évaluer la capacité des élèves à appliquer leurs connaissances théoriques et pratiques dans des contextes extrascolaires et scolaires ; remplacer progressivement les épreuves établies sur la base de niveaux par des épreuves établies sur la base de critères (en d'autres termes, mesurer le succès ou l'échec d'un élève en fonction de critères représentatifs de compétences et non pas en fonction des performances des autres candidats) ; introduire dans les notes d'examen des élèves, des informations issues d'évaluations faites par les enseignants; prendre en compte, dans les examens nationaux, les informations figurant dans les dossiers scolaires et l'évolution des performances des élèves afin de refléter l'acquisition de compétences autres que celles évaluées par un examen final écrit: et veiller à ce que les examens différencient les élèves sur la base de critères pertinents au regard de l'offre d'enseignement (Bude & Lewin, 1997 ; Kellaghan & Greaney, 1992). Toutes ces réformes semblent en soi souhaitables et devraient aboutir à des évaluations plus adaptées et plus cohérentes du niveau scolaire.

2.8. Utilisation des données tirées des examens pour refléter le niveau scolaire

46. Compte tenu de la prévalence des examens nationaux en Afrique, il apparaît raisonnable de se demander si les données ainsi obtenues donnent des indications sur les niveaux scolaires susceptibles d'être utiles aux responsables de l'élaboration des politiques et aux enseignants. On est en droit de se demander si ces données peuvent aider à déterminer si le niveau scolaire s'améliore, se maintient ou se dégrade. Nous étudierons trois sources d'information sur les performances obtenues aux examens : les rapports de jurys d'examen, d'autres mécanismes de retour d'informations et une analyse sur les catégories d'élèves en difficulté scolaire.

2.8.1. Rapports de jurys d'examen

47. L'analyse des performances obtenues par les élèves aux examens est monnaie courante. Habituellement, les résultats sont consignés dans des rapports de jury dans lesquels les examinateurs soulignent les points forts et les points faibles d'un candidat et les lacunes constatées dans l'enseignement. Ces rapports sont très variables, en ce qui concerne la qualité, le degré de détails et la pertinence des informations qu'ils transmettent aux écoles et aux enseignants. Quant à savoir si les informations qu'ils renferment sont prises en compte par les principaux intéressés, rien ne permet de tirer des conclusions à cet égard.

48. On trouvera dans les rapports de jurys d'examen au Swaziland des exemples concernant le type d'informations qu'ils renferment et qui reflètent assez bien la situation dans d'autres pays. Ces rapports sont élaborés à partir des copies rendues par les élèves à l'examen du certificat d'études primaires et à l'examen du certificat de fin de premier cycle d'études secondaires. Ils contiennent des commentaires sur les atouts et les lacunes des élèves, sur les compétences des enseignants et sur la mise en œuvre des programmes scolaires dans les écoles; ainsi que sur la façon dont les aptitudes linguistiques des élèves influent sur leurs résultats (*Encadré 1*).

Encadré 1. Rapports de jurys d'examen au Swaziland

Les rapports des examens du certificat d'études primaires ont fait état du faible niveau de vocabulaire et de compréhension en anglais des candidats. Les candidats avaient également des difficultés à construire des phrases écrites, ainsi que des lacunes en grammaire, orthographe et écriture. En lecture, l'aptitude des candidats à lire des passages de texte et à en extraire les informations pertinentes était médiocre. Le niveau d'anglais des élèves a été mesuré non seulement par rapport aux notes obtenues à l'épreuve d'anglais, mais dans d'autres matières. Ainsi, il semble qu'en sciences, sciences humaines, agriculture et économie domestique, les candidats aient été pénalisés du fait de leur incapacité à comprendre certains mots ou certaines questions.

Dans les rapports de jurys d'examen pour le certificat de fin de premier cycle d'études secondaires, la question du niveau d'anglais des élèves est de nouveau mentionnée, non seulement s'agissant des notes obtenues à l'épreuve d'anglais (difficultés des candidats à construire des phrases et à utiliser les temps corrects, mauvaise ponctuation et mauvaise orthographe, difficultés à répondre aux questions faisant appel au raisonnement, à la discussion et au symbolisme), mais aussi dans d'autres épreuves. A titre d'exemple, en mathématiques, on a constaté l'incapacité des élèves à répondre aux questions; en sciences appliquées, les candidats ont été gênés du fait de leur faible niveau d'anglais, car ils ne comprenaient pas les questions et avaient du mal à rédiger leurs réponses.

Les rapports de jurys d'examen au Swaziland n'étaient pas limités à l'analyse des résultats des candidats. Les examinateurs ont également fait part de leur sentiment sur la mise en œuvre des programmes scolaires dans les écoles et sur l'enseignement. Ainsi, ils concluent que, dans certaines écoles, le programme de sciences appliquées n'a pas été entièrement traité, de sorte que les candidats n'ont pas répondu aux questions posées en physique et chimie. De même, de nombreux domaines en métallurgie (tôlerie, conception, projection oblique) n'ont vraisemblablement pas été abordés. En sciences appliquées, il est apparu que, dans certaines écoles, les élèves avaient reçu des informations erronées. La nécessité d'élargir l'accès des élèves à des ouvrages et autres supports a été évoquée dans les commentaires relatifs à l'épreuve de géographie (Kellaghan, 2002).

49. Le rôle de la langue dans l'enseignement et, par extension, dans les examens, soulevé dans le rapport sur le Swaziland a suscité un vif intérêt ces dernières années dans les pays anglophones et, plus encore peut-être, dans les pays francophones (*Encadré 2*). Il ne fait aucun doute que de nombreux élèves sont désavantagés lorsqu'ils suivent un enseignement dans une langue qu'ils maîtrisent mal (Naumann & Wolf, 2001). D'où la question de savoir dans quelle mesure la maîtrise d'une langue explique la variance entre les notes des candidats sans que celle-ci soit imputable à leurs compétences dans la discipline concernée. Au vu des rapports de jurys d'examen pour le Swaziland, on peut supposer que ce problème a concerné plusieurs matières. Une étude réalisée au Burundi, qui compare les performances des élèves aux épreuves de lecture, mathématiques,

sciences et agriculture organisées en français et en kirundi, est également révélatrice. Les élèves ont obtenu de meilleurs résultats en kirundi (sauf en mathématiques). De plus, les meilleurs élèves ont eu des difficultés dans les épreuves subies en langue française (Eisemon, 1990).

Encadré 2. Langue d'enseignement

La question de la langue ne se pose pas exclusivement pour les examens, mais concerne aussi l'emploi d'une langue de grande diffusion (LWC) pour l'enseignement, notamment dans des matières qui ont une grande importance pratique pour la majorité des élèves en fin de primaire qui ne poursuivent pas leur scolarité. La plupart des langues de la communauté africaine ne sont pas utilisées comme langues d'enseignement dans de nombreux pays, et, dans certains pays, pas du tout (Ouane, 2003). C'est là une question importante puisque, pour le projet MLA ("Monitoring Learning Achievement"), les pays où l'enseignement était dispensé dans la langue maternelle des élèves ont eu des notes nettement supérieures aux autres, et ce dans la majorité des matières (Chinapah et al, 2000).

Conscients du fait que la méconnaissance de la langue d'enseignement par les élèves peut être un obstacle à leur apprentissage, plusieurs pays ont mis en place des programmes expérimentaux dans lesquels l'enseignement est prodigué dans la langue maternelle des élèves (Mail, Niger et Nigeria par ex.). Les premiers résultats semblent indiquer que, chez les élèves ayant bénéficié de ces programmes, les taux de redoublement et d'abandon sont plus faibles, les performances aux examens de fin de primaire meilleures et les taux de passage dans la classe supérieure plus élevés (ADEA, 1999; Bamgbose, 1991). Toutefois, comme les effectifs des classes dans les écoles expérimentales étudiées au Niger étaient beaucoup plus faibles (22-25:1, contre 50-90:1 dans les classes traditionnelles), il n'est pas possible d'imputer les avantages incontestables de ce système à la seule langue d'enseignement.

Les raisons de l'emploi d'une langue LWC sont complexes. Elles peuvent être liées à une volonté de promouvoir l'intégration nationale, de réduire les coûts et de faciliter la communication internationale (Clayton, 1998). Une langue LWC peut aussi présenter des avantages pratiques dans des contextes hétérogènes au plan linguistique.

2.8.2. Autres mécanismes de retour d'informations

50. Outre les rapports de jurys d'examen, plusieurs pays ont mis au point des systèmes d'analyse des informations sur les performances obtenues par les élèves aux examens de fin de scolarité primaire, ainsi que de retour de ces informations aux écoles. Une réforme des examens au Kenya a donné matière à des « orientations » élaborées à partir d'une analyse des performances des candidats sur différents aspects contrôlés lors de l'examen (Kyalo, 1997; Rharade, 1997 ; Somerset, 1987, 1988, 1996). Des conseils concrets ont également été fournis dans un bulletin d'information diffusé à toutes les écoles et dans le cadre d'ateliers, de séminaires et de visites d'écoles (notamment les écoles ayant enregistré des résultats médiocres). Le Conseil des examens nationaux du Kenya continue de rédiger un bulletin, mais ne peut, faute de moyens financiers, l'adresser à toutes les écoles. Celles-ci peuvent toutefois l'acheter auprès du Conseil (J.M. Mwachihi, communication personnelle, mai 2003).

51. Le retour d'informations a également été employé dans d'autres systèmes d'examen. En Zambie, chaque école reçoit un exemplaire du rapport de jurys d'examen. Au Lesotho, des statistiques détaillées sur les performances des candidats aux examens

ont été envoyées à toutes les écoles (Kellaghan & Greaney, 1992). En Ouganda, des bulletins d'informations ont été adressés aux écoles ayant enregistré des performances médiocres aux examens, accompagnés de suggestions sur les stratégies pédagogiques susceptibles de résoudre les problèmes constatés.

2.8.3. Repérage de catégories d'élèves en difficulté scolaire

52. Les résultats d'examens peuvent permettre de repérer des différences de performances entre les garçons et les filles, d'une région à l'autre, d'une zone à l'autre (urbaine-rurale) et d'une école à l'autre. Le cas d'écoles où les élèves ont pratiquement tous réussi les examens et d'autres écoles où ils ont pratiquement tous échoué n'est pas rare. Les raisons de tels écarts sont complexes et font intervenir des facteurs qui sont liés au contexte socioculturel et aux antécédents des élèves et, partant, peu sensibles aux effets d'une politique. Néanmoins, l'octroi par les autorités éducatives de conseils, de soutien et de ressources à des écoles ou des régions qui enregistrent de mauvais résultats peut être source d'amélioration. Les analyses mettant en évidence les écarts entre garçons et filles ou entre régions peuvent aussi servir de base à une réforme des procédures d'examen et être un moyen de s'assurer qu'elles ne défavorisent pas certains sous-groupes de candidats.

2.8.4. Limites des données fournies par les examens

53. Si l'analyse des résultats d'examen donne des indications sur le niveau scolaire des élèves dans le système éducatif et permet de repérer les écoles où les élèves ont des résultats médiocres et qui méritent donc une plus grande attention, elle n'en comporte pas moins des limites. L'une d'elles est que les examens nationaux ne révèlent en général le niveau des élèves que dans certaines matières du programme seulement. En effet, le nombre de matières présentées à l'examen est restreint. Par ailleurs, même dans ces matières, les épreuves mettent l'accent sur le contenu du programme scolaire et sur des compétences connexes destinées à séparer nettement les élèves admis et les élèves recalés. Ainsi, les résultats des élèves les plus faibles ne sont pas toujours clairement représentés.

54. Les résultats d'examens sont parfois utilisés pour suivre l'évolution du niveau scolaire des élèves dans le système éducatif au fil des années afin de déterminer si le niveau scolaire augmente, baisse ou se maintient. Mais une telle analyse ne va pas sans difficulté. Tout d'abord, l'examen diffère d'une année sur l'autre et il n'est donc pas possible de savoir si le degré de difficulté reste constant avec le temps. Deuxièmement, les résultats d'examens sont souvent normalisés et établis sur la base de niveaux et ils reflètent leur fonction sélective, de sorte que la moyenne annoncée est fixe d'une année sur l'autre, ce qui ne montre pas de changement dans les niveaux. Troisièmement, comme l'offre éducative continue de s'étendre (conformément aux objectifs escomptés de l'EPT) et que le nombre de candidats aux examens nationaux augmente, les caractéristiques de ces candidats évoluent. D'où une incidence potentielle sur le niveau scolaire moyen des candidats, sans toutefois qu'une comparaison des résultats d'examens d'une année sur l'autre ne permette de savoir dans quelle proportion (Greaney & Kellaghan, 1996a; Kellaghan, 1996a).

2.9. Utilisation des examens pour améliorer le niveau scolaire des élèves

55. L'analyse des résultats d'examen peut fournir des informations, certes limitées, mais néanmoins utiles, sur le niveau scolaire des élèves dans le système éducatif. Cependant, le rôle des examens est beaucoup plus radical et proactif lorsqu'il s'agit de les utiliser, non pas seulement comme révélateur de l'état des systèmes éducatifs, mais comme levier d'une réforme. Ce genre de propositions a suscité un grand intérêt notamment du fait que, *lorsque les enjeux liés aux performances obtenues aux examens étaient élevés* (c'est-à-dire quand les performances obtenues aux examens sont déterminantes pour les élèves et souvent aussi pour les enseignants), on constate qu'un alignement s'opère entre ce qui est enseigné à l'école et les objectifs de l'examen. L'idée d'un « enseignement induit par l'évaluation » a ainsi été défendue aux Etats-Unis par Popham (1983, 1987); à ses yeux, la modification d'une évaluation ne doit pas nécessairement refléter une modification du programme scolaire. En réalité, la modification d'une évaluation doit être première dans le processus et elle est censée dicter le programme enseigné dans les écoles. Pour Popham, l'important était que « les tests de niveau dont on espère qu'ils répondent à des finalités pédagogiques contiennent un nombre raisonnable d'objectifs d'évaluation, de telle sorte que les enseignants puissent articuler leur mode d'enseignement autour d'un nombre raisonnable d'objectifs pédagogiques » (p. 25). L'approche, comme on le verra, va à l'encontre de l'idée plus traditionnelle selon laquelle le contenu d'une évaluation doit refléter le programme; à titre d'exemple, cette approche a servi de base aux réformes de l'examen national et du système d'évaluation en Namibie, réformes jugées nécessaires du fait de la réforme du programme scolaire (ADEA, 1999).

56. S'agissant de l'éducation dans des pays en développement, plusieurs commentateurs ont évoqué l'intérêt d'un enseignement induit par l'évaluation comme moyen de modifier les pratiques à l'échelon de la salle de classe, quand des enjeux liés aux performances sont élevés. Le principe était relativement simple : s'assurer que les examens couvrent un contenu vaste et pertinent et évaluer les connaissances théoriques et pratiques essentielles; les enseignants ajusteront alors leur enseignement en conséquence. Dans le même ordre d'idées, Eisemon et al (1987) considèrent qu'améliorer l'enseignement via une amélioration des examens est foncièrement une bonne stratégie, tandis que, pour Little (1982), « l'amélioration des examens pourrait contribuer à faire évoluer le système éducatif vers un système qui encourage, au lieu de dévaloriser, les résultats souhaitables » (p. 228). De surcroît, modifier les examens est un moyen, non seulement efficace, mais aussi non coûteux, d'influer sur la qualité de l'enseignement et de l'apprentissage dans les écoles (Heyneman & Ransom, 1992).

2.10. Effets des modifications apportées aux examens

57. Peut-on imaginer qu'une modification des examens nationaux améliorera la qualité d'apprentissage des élèves ? Nous tenterons de répondre à cette question en étudiant les données disponibles par rapport à trois questions plus spécifiques relatives aux objectifs des réformes. Une modification du contenu contrôlé par l'examen apporte-t-elle une modification du contenu enseigné en classe aux élèves ? Une modification des examens peut-elle induire une amélioration du niveau scolaire des élèves ? Et, une modification des examens peut-elle se traduire par une modification des facultés cognitives des élèves ?

58. Une modification du contenu contrôlé par l'examen apporte-t-elle une modification du contenu dispensé en classe aux élèves ? Dans maints pays, de nombreux indices tendent à montrer qu'une modification des examens influe sur le contenu dispensé aux élèves pendant les cours - les matières qui retiennent l'attention et les sujets traités au sein des matières enseignées. Les exemples abondent : en Australie, à la suite d'une modification des examens de chimie; en Irlande, pour les examens de sciences physiques, mathématiques et chimie; en Belgique, pour les examens de primaire; en langues vivantes pour les New York Regents examination; et en mathématiques pour les tests de niveau du College Entrance Examination Board (Madaus & Kellaghan, 1992).

59. Au Kenya, à la suite de l'introduction du kiswahili et de disciplines pratiques pour le Certificat d'études primaires du Kenya, ces disciplines ont été beaucoup plus largement traitées dans les écoles malgré les difficultés rencontrées au plan des installations, des manuels et des compétences des enseignants (Eisemon, 1990). Dans les années 80, pour le Common Entrance Examination (en fin de scolarité primaire) organisé à Trinidad et Tobago, les questionnaires à choix multiple portant sur le style et la structure littéraires ont été remplacés par la rédaction d'un essai, entraînant une augmentation du nombre de devoirs écrits à rendre sur une année par les élèves afin qu'ils puissent se familiariser avec la technique argumentaire et apprennent à appliquer leurs connaissances pour résoudre des questions et des problèmes nouveaux. London (1997) indique, sans toutefois fournir de preuves à l'appui de cette affirmation, que « la rédaction d'essais est désormais enseignée dans les écoles depuis près de dix ans... [et que] la majorité des enseignants... se déclarent satisfaits ... que, dans l'emploi du temps quotidien de la classe, un nombre d'heures raisonnable soit imparti... à la rédaction d'essais » (p. 144).

2.10.1. Une modification des examens peut-elle induire une amélioration du niveau scolaire des élèves ?

60. Que ce soit pour appuyer la thèse selon laquelle une modification des examens induit une amélioration du niveau scolaire des élèves ou au contraire pour la réfuter, les preuves empiriques disponibles sont rares.

61. L'étude la plus fréquemment citée en la matière est celle réalisée par Somerset (1987, 1988) qui concerne les effets produits par la réforme de l'examen de fin du cycle d'éducation de base, réforme qui a été instaurée dans les années 70 au Kenya. Les objectifs de cette réforme étaient (i) d'améliorer l'efficacité de l'examen en tant qu'instrument de sélection; (ii) de donner aux élèves moins favorisés (milieu rural, familles à faible revenu, filles) une chance plus grande de montrer leurs aptitudes et, par là même, d'avoir accès à l'enseignement secondaire; (iii) d'encourager et d'aider les enseignants à transmettre à tous les élèves (notamment à ceux dont l'éducation de base semble la seule et ultime éducation dont ils pourront bénéficier) un ensemble plus cohérent de compétences cognitives; (iv) de relever le niveau général de qualité de l'éducation primaire; et (v) de réduire les écarts d'un district à l'autre et d'une école à l'autre. A cette fin, une modification substantielle du contenu des examens a été opérée, un tableau d'honneur des écoles a été publié à titre d'incitation pour les établissements scolaires; et un bulletin d'informations sur les résultats obtenus par les élèves aux examens a été diffusé.

62. Malheureusement, faute de données, il n'est pas possible de déterminer si ces réformes ont abouti à une amélioration globale des performances des élèves, du fait que les résultats afférents à chaque année sont étalonnés (par rapport à une moyenne de 50 et un écart-type de 15). On peut également supposer qu'une comparaison des résultats bruts ne serait pas plus utile. Les données concernant le pourcentage de « reçus » auraient pu

servir de base à une comparaison des performances d'une année sur l'autre, à condition d'avoir la certitude que les critères restent identiques. Par contre, les données disponibles sur les performances ont permis de comparer l'évolution des performances relatives des candidats avec le temps, et ce dans différents districts. Si, dans les premiers temps, les réformes ont eu pour effet de creuser les écarts de niveau scolaire entre les districts, cette tendance s'est inversée au bout de quatre ans d'application du système : dans la quasi-totalité des districts où les performances étaient relativement médiocres, une amélioration spectaculaire de ces performances par rapport à d'autres districts a été enregistrée (Somerset, 1987).

2.10.2. Une modification des examens peut-elle se traduire par une modification des facultés cognitives des élèves ?

63. Un objectif majeur des partisans des réformes est d'utiliser les examens pour développer des facultés cognitives supérieures chez les élèves. L'un des inconvénients des examens dans leur forme actuelle est que, dans la plupart des cas, ils n'évaluent que les aptitudes de base (mémorisation et reconnaissance), de sorte que l'enseignement scolaire est axé sur ces aptitudes de base. Par contre, si les examens exigent des élèves qu'ils fassent preuve de compétences de plus haut niveau, alors d'aucuns considèrent que les enseignants adapteront le contenu de leur enseignement et leurs méthodes pédagogiques pour répondre à cette exigence.

64. Les données relatives à l'effet des examens sur l'enseignement et sur l'acquisition de compétences de plus haut niveau par les élèves sont hétérogènes. Si l'on manque de données sur l'expérience réalisée au Kenya en la matière, on constate que Eisemon et al (1987) sont incapables de déterminer si, à la suite de l'introduction des nouveaux types de questions, une place plus importante a été donnée dans l'enseignement primaire à la résolution de problèmes, au raisonnement et à l'explication; ceci laisse supposer qu'il n'y a pas lieu d'escompter de grands changements au niveau des facultés cognitives des élèves.

65. D'autres éléments montrent que les enseignants ne réagissent pas toujours aux exigences fixées par les examens, d'où des répercussions sur les compétences que les élèves acquièrent. Une étude menée au Swaziland (Rollnick et al, 1998) révèle que les stratégies pédagogiques employées par les enseignants en chimie étaient liées au type de questions posées dans deux examens. Pour le premier examen ("Junior Certificate Examination", en fin du premier cycle d'études secondaires), les questions faisaient en général appel à la mémorisation et au savoir, tandis que, pour l'examen suivant ("O-level", brevet en fin de second cycle d'études secondaires), trois niveaux cognitifs (mémorisation, connaissances et compréhension, analyse des données) étaient évalués. Or, l'enseignement préparatoire à ces deux examens était foncièrement identique. Dans l'un et l'autre cas, les enseignants étaient essentiellement attentifs à l'acquisition d'un savoir factuel et recourraient donc, dans leurs cours et les contrôles, à des questions cognitives de base. Comme les enseignants au niveau du brevet ont appliqué les mêmes stratégies que leurs collègues de premier cycle secondaire, c'est-à-dire sans prévoir d'enseignement ni d'épreuves sur les facultés cognitives de plus haut niveau utiles pour l'examen, il n'est pas surprenant qu'un tiers des élèves ait échoué en sciences (physiques, chimie) à l'épreuve de brevet 2001 (Kellaghan, 2002). Toutefois, d'après Rollnick et al, on peut penser qu'une formation continue pourrait faire évoluer les stratégies adoptées par les enseignants.

66. Il apparaît par ailleurs que, lorsque les enseignants bénéficient de conseils pour préparer leurs élèves aux examens, ainsi que d'explications, dans le cadre d'une

formation continue, pour comprendre les exigences imposées dans les examens, l'introduction dans les épreuves de tâches nécessitant que les candidats fassent preuve de capacités de plus haut niveau peut conduire les enseignants à mettre davantage l'accent sur cet aspect dans leurs cours. Dans une étude menée sur des écoles primaires de Nairobi de niveau 8 ("Standard 8"), les enseignants ont été invités à préparer les élèves à deux modèles de questions au titre d'un examen blanc. Un premier modèle a été préparé par le Conseil national des examens du Kenya; le second a été spécialement conçu pour faire ressortir des facultés cognitives de niveau plus élaboré. Dans ce dernier cas, il ne s'agissait pas pour les élèves de trouver la bonne réponse, mais de tirer des conclusions. Ce dernier modèle a conduit les enseignants à modifier leur façon de préparer les élèves à l'examen. En outre, les élèves qui avaient été préparés pour ce type d'épreuve ont obtenu de meilleurs résultats aux deux examens que ceux qui n'avaient pas reçu cette préparation spéciale (Eisemon, 1990).

2.10.3. Mécanismes destinés à modifier les pratiques scolaires et les résultats des élèves

67. Modifier un examen, le contenu de l'épreuve et le type de facultés cognitives requises des candidats peut en soi suffire à faire évoluer les pratiques scolaires et les connaissances théoriques et pratiques des élèves. Il semble que ce soit le cas plus particulièrement lorsque l'on modifie les matières figurant au programme de l'examen (par ex. lorsqu'une épreuve pratique est introduite ou que les élèves doivent rédiger un essai et répondre à des questions nécessitant des réponses courtes). En l'occurrence, on peut s'attendre à ce que les enseignants s'adaptent au programme et aux exigences de l'examen et à ce qu'ils donnent aux élèves les moyens nécessaires pour réussir.

68. Mais il peut arriver que des enseignants aient besoin d'être davantage incités à changer leurs habitudes ou encore que des enseignants ne soient pas suffisamment préparés pour assumer un tel changement. Des enseignants peuvent par exemple se sentir incompetents pour enseigner tel ou tel aspect pratique d'une discipline scientifique ou incapables de changer leur méthode d'enseignement pour favoriser le développement de capacités de réflexion plus élaborées. Dans ce cas, les concepteurs des réformes « en matière d'examens et les autorités scolaires ont eu recours à différents moyens pour induire un changement de comportement chez les enseignants, parmi lesquels : instaurer ou renforcer la compétition entre les écoles, fournir des informations sur les écoles, octroyer des conseils ou des ressources supplémentaires

2.10.4. Compétition entre écoles

69. Certains pays ont tenté d'influer sur le comportement des enseignants en publiant les résultats d'examen des élèves issus de différentes écoles. La publication d'un « palmarès » des résultats, comportant les statistiques de performances moyennes de chaque école et de chaque district, a ainsi permis à chaque école de se situer par rapport aux autres écoles d'un même district et à chaque district de comparer son classement par rapport à d'autres districts. Elle a constitué l'un des axes majeurs de la réforme des examens au Kenya et les informations ainsi obtenues ont été qualifiées « d'informations incitatives » (Somerset, 1987). L'idée de base de cette méthode était que la diffusion des résultats des écoles créerait un climat de compétition entre elles et, par voie de conséquence, inciterait les enseignants à améliorer leurs pratiques pédagogiques (Chapman & Synder, 2000). De plus, ces informations étaient destinées à donner naissance à une sorte de « marché » des écoles où les parents qui avaient le choix entre plusieurs établissements feraient leur « marché » avec « preuves » à l'appui sur les performances des uns et des autres.

70. Il n'y a aucun doute que la publication d'un palmarès des résultats des évaluations désignant nommément les écoles performantes et les écoles moins méritantes peut avoir un impact sur le comportement de ces écoles. Durant les années 90 au Sénégal, un système de gestion utilitaire avec diffusion des performances des écoles dans la presse a été institué. Entre 1995 et 1998, le pourcentage de réussite à l'examen de fin d'études primaires est alors passé de 30 % à 48 %. Quant au taux de scolarisation des filles, il a progressé de 40 % à 77 % (ADEA, 2002). Pour autant, on ne saurait imputer ces progrès au seul fait de la publication des résultats, car d'autres réformes ont été mises en œuvre dans le même temps (élaboration de définitions de poste; augmentation des inspections dans les écoles; organisation de séminaires et de journées portes ouvertes).

71. Malgré la diffusion des palmarès des écoles au Kenya (qui n'a plus cours aujourd'hui), au Sénégal et dans d'autres pays, plusieurs facteurs indiquent que cette pratique est plus complexe qu'il n'y paraît à première vue, qu'elle n'est pas sans soulever des difficultés, tant s'en faut, et qu'elle risque même d'être trompeuse (Encadré 3). Hormis les points évoqués dans l'Encadré 3, il ne faut pas oublier lorsque l'on analyse la contribution apportée par l'école au développement des élèves que les résultats de ces derniers dépendent d'une multitude de paramètres extérieurs à l'école. S'agissant de l'incidence de l'école sur les performances des élèves, il importe de considérer également les facteurs indépendants du cadre purement scolaire. Il semble par exemple judicieux de se demander si les élèves ont de mauvais résultats parce qu'ils ne sont pas motivés ou n'ont pas bénéficié de soutien dans leur famille, parce que l'école manque de moyens ou encore parce que les enseignants sont médiocres. Qui plus est, à supposer que le niveau des élèves à leur arrivée dans une école diffère d'une école à l'autre, il serait injuste et trompeur d'utiliser ultérieurement un critère de mesure qui ne tient pas compte de cet écart initial : en effet, ce critère ne traduira pas la capacité éventuelle d'une école à faire progresser ses élèves jusqu'au niveau reflété dans les résultats obtenus à un examen public (Kellaghan & Greaney, 2001b).

2.10.5. Octroi de conseils aux écoles

72. Le but est en l'occurrence d'analyser les résultats d'examen des élèves et d'en communiquer le bilan aux écoles concernées. Comme on l'a mentionné précédemment, ces données peuvent être consignées dans des rapports de jurys d'examen et, plus récemment, dans des bulletins d'informations adressés aux écoles. C'est ainsi que l'un des axes prioritaires du bulletin d'informations élaboré dans le cadre de la réforme des examens au Kenya a consisté à fournir des « informations et conseils » mettant en évidence les principaux domaines d'études et de compétences dans lesquels les élèves avaient des difficultés et contenant des suggestions à l'adresse des enseignants pour résoudre ces difficultés (Somerset, 1987, 1988). Selon Somerset, l'aspect compétitif n'a pas été suffisamment souligné dans les réformes; de même, il est nécessaire d'apporter un soutien et des conseils aux enseignants afin qu'ils puissent réorienter leurs efforts.

Encadré 3. Inconvénients des palmarès

- Les performances des écoles (et donc leur classement par rapport à d'autres écoles) peuvent varier selon le type de résultats pris en compte. Ainsi, le classement des écoles sera différent selon qu'il est établi sur la base des résultats obtenus à un examen public, des performances dans des matières de base du programme, d'une évaluation des inspecteurs d'établissement ou d'une appréciation globale de l'efficacité de l'école.
- Les classements opérés sur la base de la même évaluation peuvent eux-mêmes varier en fonction du critère de « réussite » appliqué. Ainsi, un classement établi par rapport au pourcentage d'étudiants ayant obtenu de très bons résultats (mentions) peut être différent d'un classement établi par rapport au pourcentage d'étudiants ayant obtenu des résultats moins bons, mais satisfaisants (pourcentage d'admis par ex.).
- Compte tenu du caractère imprécis des instruments d'évaluation, même soigneusement conçus, un léger écart entre les notes respectives des écoles pourra être le fait du hasard. En revanche, un écart de quelques points dans une note pourra, dans un système scolaire très étendu, entraîner une très grande différence de classement entre les écoles.
- Quel que soit le critère de mesure retenu, les résultats sont variables d'une année sur l'autre. En Grande Bretagne par exemple, parmi les 250 premières écoles, le classement d'une école peut varier de 123 places d'une année à l'autre. Cette variation peut être imputable à des différences liées aux instruments de mesure, aux promotions d'élèves ou aux enseignants.
- La compétition est inopérante lorsqu'une région ne compte qu'une seule école.

2.10.6. Octroi de soutien à des écoles peu performantes

73. Outre les informations sur les performances, une autre solution consiste à repérer les écoles « peu performantes », à en rechercher les raisons et à leur fournir une aide appropriée. Ce système suppose d'expliquer en détails la signification des résultats (en se référant par exemple au niveau exigé et au type d'erreurs commises par les élèves), puis de fournir des conseils pratiques sur les compétences à développer, les séquences à mettre en œuvre et les types d'exercice à donner aux élèves. La pertinence de ces conseils par rapport aux objectifs poursuivis doit être visible en termes concrets. Ce système sera plus efficace si les enseignants peuvent espérer une « récompense » directe en retour de leurs efforts sous la forme d'une amélioration des performances à l'examen suivant (Grisay & Mählch, 1991).

74. S'il apparaît que ces écoles sont peu performantes en raison du manque de manuels scolaires, la solution du problème peut être relativement simple. Mais, en général, les causes sont plus complexes (Chapman & Snyder, 2000) et nécessitent par là même des actions plus complexes. Ainsi au Kenya, les bureaux éducatifs de district ont fourni aux écoles des conseils d'ordre pédagogique ou autre pour les aider à résoudre les difficultés mises en évidence dans les résultats (Somerset, 1988).

75. Ceci dit, même lorsque les moyens existent, il peut arriver que des enseignants ne sachent pas comment les utiliser pour améliorer le niveau des élèves. En effet, lorsqu'ils manquaient de moyens, ces enseignants ont su s'adapter à la situation; ceci

ne veut pas dire que, dès lors qu'ils ont les moyens, ils vont automatiquement changer leur méthode d'enseignement et en faire le meilleur usage.

76. Un autre problème lié aux écoles peu performantes est qu'elles sont souvent implantées dans des régions particulièrement difficiles d'accès et qu'elles desservent des populations minoritaires qui n'ont guère de poids au regard de la politique nationale (Chapman & Snyder, 2000).

2.10.7. Liens avec les programmes scolaires

77. A une échelle plus large, les autorités responsables des examens, les inspecteurs du ministère de l'éducation et le personnel chargé des programmes scolaires peuvent analyser les résultats et leurs répercussions sur l'élaboration et l'application des programmes, sur l'enseignement et sur la préparation des manuels scolaires, comme cela s'est produit au Kenya (Somerset, 1988). Cette pratique témoigne de l'intérêt porté à l'instauration de relations étroites entre les autorités responsables des examens et les autorités responsables des programmes scolaires (Eisemon, 1990).

2.10.8. Problèmes particuliers en cas d'enjeux importants liés aux résultats des évaluations

78. Comme on l'a vu précédemment, les partisans de réformes des examens considèrent que, si le contenu des épreuves est plus large et que l'échelon taxinomique des facultés cognitives requises pour répondre aux questions posées est plus élevé, les effets devraient s'en ressentir au niveau de la méthode pédagogique et du contenu de l'enseignement dispensé aux élèves à l'école. En écho à cette suggestion, Somerset (1996) indique que « il n'y a pas de raison en soi pour que les examens aient des effets secondaires indésirables sur l'enseignement et l'apprentissage; ces effets indésirables n'apparaissent que si les examens sont de mauvaise qualité » (p. 282). Tout en reconnaissant que l'incidence des examens sur les pratiques scolaires est due aux enjeux importants liés aux performances, cette théorie ne tient pas compte d'un autre élément qui a été constaté depuis longtemps et selon lequel, quelle que soit la qualité des examens, il s'ensuivra des effets négatifs, s'ils ne sont pas intentionnels, si des sanctions sont liées aux performances (Kellaghan & Greaney, 1992; Madaus & Greaney, 1985; Madaus & Kellaghan, 1992). Ces effets sont évoqués dans le présent chapitre (voir *Encadré 4*).

Encadré 4. Incidence des examens sur les enseignants

L'incidence des examens sur l'enseignement transparaît clairement dans la description du comportement assez typique des enseignants au Burundi. Au niveau des caractéristiques de l'enseignement, on a par exemple relevé : une augmentation du nombre d'heures d'enseignement par rapport à celui fixé pour les écoles primaires (les samedis après-midi par ex.); la révision régulière des aspects considérés comme essentiels; des contrôles de connaissance fréquents pour tester les éventuelles lacunes et améliorer les compétences exigées à l'examen; un diagnostic des causes d'erreurs commises par les élèves en recourant à des méthodes pédagogiques diversifiées; et l'utilisation des périodes scolaires spécialement consacrées à l'étude de domaines de formation préprofessionnelle non présentés à l'examen (l'agriculture par ex.) pour enseigner les matières présentées à l'examen (Eisemon, 1990). Des pratiques identiques ont été constatées dans d'autres pays où les résultats d'examen étaient déterminants pour les études et pour l'avenir des élèves ou étaient utilisés pour évaluer les enseignants (Kellaghan & Greaney, 1992).

79. Si l'alignement de l'enseignement sur les examens, selon lequel les enseignants répercutent les principales exigences des examens dans leur enseignement, est censé être l'un des avantages des examens comportant un enjeu important, le revers de la médaille est que seul un sous-ensemble ou un échantillon d'un domaine d'études plus vaste sera évalué dans le cadre de cet examen. Si toute l'attention est concentrée sur l'examen lui-même, les écoles n'accorderont pas aux matières non présentées à cet examen tout l'intérêt qu'elles méritent au titre du fait qu'elles figurent au programme scolaire. De plus, dans ces matières elles-mêmes, les enseignants auront tendance à insister sur l'enseignement et le contenu de l'épreuve et donc à ne traiter qu'une certaine partie du programme; à terme, l'examen finira par réduire et délimiter le champ du programme scolaire traité (Le Mahieu & Leinhardt, 1985).

80. Un autre inconvénient des examens comportant un enjeu élevé est qu'ils influent souvent sur les stratégies d'enseignement, sur les stratégies d'apprentissage, sur l'implication des élèves dans leur apprentissage et sur leur attitude vis-à-vis de leur apprentissage. Les enseignants vont s'appesantir davantage sur les exercices d'application et exiger des élèves qu'ils adoptent des stratégies d'apprentissage superficielles ou à court terme (telles que mémorisation, répétition, savoir par cœur). Les élèves vont adapter leur comportement en conséquence. Ils vont aussi être plus sensibles aux comparaisons sociales (être meilleur que les autres), attitude qui est renforcée par la pratique courante d'applaudir les élèves qui donnent une bonne réponse en classe. Ils vont avoir plus de difficultés à acquérir des compétences transférables d'un niveau plus élaboré; les élèves vont considérer les tâches d'apprentissage comme peu intéressantes en-soi et, pour peu qu'il n'y ait pas de récompense à attendre, ils seront peu enclins à faire des efforts (Kellaghan, Madaus, & Raczek, 1996). Le risque est par conséquent que certains objectifs majeurs de la réforme de l'éducation ne soient pas atteints.

81. Autre problème lié aux examens comportant un enjeu élevé : les efforts et le temps consacrés par les enseignants et par les élèves à la préparation de l'examen sont considérables. Outre les heures d'enseignement perdues (à cause des « examens blancs » par exemple). Cela peut représenter presque une semaine par mois pour les élèves du Mali et de la Guinée. Les exercices de préparation risquent de fausser l'enseignement et l'apprentissage. Par exemple, si l'examen comporte des questionnaires à choix multiple, l'enseignement normal risque d'évoluer dans le même sens, avec des affirmations suivies d'une série d'options parmi lesquelles l'élève doit choisir sa réponse. La place importante donnée aux compétences requises pour l'examen peut aussi inciter les élèves à axer leurs efforts sur la maîtrise de soi pendant l'examen, plutôt qu'à développer une maîtrise de la matière présentée à l'examen et à affiner des compétences durables.

82. Autre sujet de préoccupation relatif aux examens comportant un enjeu élevé : les efforts des élèves risquent de ne pas être limités à l'expérience éducative acquise dans le niveau de classe sanctionné par l'examen, même si ces efforts sont plus intenses dans la réalité. Il est probable que, dans des classes de niveau moins élevé, les matières figurant au programme de l'examen recevront davantage d'attention, et ce au détriment d'autres objectifs du programme scolaire. La forme même des épreuves peut avoir une incidence sur l'enseignement. Par exemple, les questionnaires à choix multiple ont été utilisés non seulement en classe, mais aussi dans la méthodologie appliquée par les enseignants durant les premières années de scolarité primaire.

83. Plus généralement, les examens comportant un enjeu élevé ne permettent pas de tirer des conclusions générales quant à une amélioration potentielle des résultats des élèves. Il n'y a pas lieu de s'étonner que, si les enseignants concentrent leurs efforts sur le contenu présenté à l'examen, les élèves obtiennent de meilleures notes à l'examen en question. Cependant, l'amélioration est plus visible dans les domaines où les objectifs des programmes scolaires se chevauchent avec le contenu de l'épreuve (Le Mahieu &

Leinhardt, 1985). Dans ce cas, on assiste à un déplacement des objectifs : la préparation de l'examen détourne l'attention des objectifs du programme scolaire au profit de la nécessité de satisfaire aux exigences de l'examen. Par conséquent, les conclusions que l'on peut tirer des résultats obtenus à l'examen risquent d'être erronées. On ne peut pas conclure que de bons résultats ou de meilleures notes sont la preuve d'un bon niveau sur l'ensemble du programme scolaire (ce qu'un examen est censé évaluer). L'examen n'est alors pas un moyen fiable d'évaluer un niveau global (voir aussi Le Mahieu, 1985 ; Linn, 1983, 2000).

84. D'une façon générale également, l'existence d'un enjeu important peut être propice aux pratiques malhonnêtes. Soucieux d'obtenir de bons résultats, les élèves (et parfois aussi les enseignants ou d'autres) peuvent être tentés de recourir à différentes formes de tricherie afin d'être avantagés par rapport aux autres candidats (Greaney & Kellaghan, 1996b). Ces pratiques malhonnêtes revêtent des formes diverses : copier sur un autre élève pendant une épreuve, complicité entre étudiants et surveillants d'examen, introduction de documents interdits dans la salle d'examen, corruption ou manœuvres d'intimidation auprès des examinateurs et achat des sujets d'examen (voir Encadré 5). La pression en faveur de telles pratiques peut être forte. On a signalé le cas de surveillants qui ont été tués parce qu'ils refusaient de coopérer. On ignore l'ampleur de la fraude aux examens, mais on peut supposer qu'elle est considérable. En Ouganda dernièrement, à l'examen du certificat d'études les résultats de 3,16 % des candidats ont été annulés pour fraude. Etant donné que ces pratiques portent atteinte à la crédibilité d'un système d'examens, divers moyens ont été mis en œuvre dans de nombreux pays pour les éviter (faire imprimer les sujets d'examen dans un autre pays ou isoler les correcteurs pendant la période de correction), les détecter (comparer les réponses d'élèves placés côte à côte au cours d'un examen par ex.) et les sanctionner (infliger une peine d'emprisonnement par ex.).

Encadré 5. Quelques exemples de fraudes à l'examen

Avant l'examen

Fuites avec divulgation du sujet d'examen par des responsables, auteurs des sujets, modérateurs ou administrateurs scolaires; affectation intentionnelle de candidats à certains centres par des responsables.

Pendant l'examen

Usurpation d'identité; aide extérieure (via des assistants, des téléphones cellulaires par ex.); introduction frauduleuse de documents (dans les vêtements par ex.), copie et collusion entre candidats; manœuvres d'intimidation auprès de surveillants (par des candidats, des assistants extérieurs, des organismes publics, des responsables politiques, des journalistes et des syndicats d'enseignants par ex.); substitution de sujets; utilisation de centres « fantômes » (qui n'existent pas).

Après l'examen

Substitution de sujets; manœuvres d'intimidation auprès de correcteurs avec le soutien de fonctionnaires corrompus; collusion entre le candidat et le correcteur (candidats qui donnent l'adresse de leur domicile par ex.); falsification de fichiers et de feuilles de résultats; délivrance de faux diplômes.

85. Enfin, il convient d'être attentif au fait que les examens comportant un enjeu élevé peuvent être contradictoires avec certains aspects de la politique d'éducation. Le cas peut se produire lorsque des enseignants, du fait que leur réputation dépend des résultats de leurs élèves aux examens, s'intéressent essentiellement aux élèves qui ont le plus de

chances de réussir. Il est alors impossible d'espérer atteindre l'objectif d'une Education pour tous et offrir à tous les élèves une éducation de base complète et de bonne qualité.

86. Privilégier un taux de réussite élevé aux examens peut également conduire à un taux élevé de redoublement et d'abandon, associés à des dépenses élevées et à un taux de rendement interne extrêmement bas (voir Eimeson, 1997 ; N'tchougan-Sonou, 2001). Au Kenya, l'une des raisons du faible taux de passage de la sixième à la septième année tient au fait que les écoles ont découragé les élèves de passer l'examen du Certificat d'études primaires du Kenya par crainte d'être mal classées au palmarès public (Ackers et al, 2001). Un autre risque est de favoriser le redoublement, l'idée étant que le redoublement améliore les chances d'un élève de réussir aux examens et donc d'entrer dans le cycle secondaire. Le redoublement n'est naturellement pas un phénomène propre au Kenya; c'est aussi une caractéristique des pays francophones (Burkina Faso, Burundi, Cameroun, Tchad, Comores, Guinée, Guinée-Bissau, Niger, Rwanda, Sénégal). Les examens comportant un enjeu élevé peuvent faire obstacle aux politiques de passage automatique menées par certains pays (qui d'ailleurs existent au Kenya).

2.10.9. Orientations pour l'amélioration des examens nationaux

87. Le fait que des examens nationaux comportent des enjeux importants liés aux performances présentera toujours des inconvénients. Il est peu probable que leur rôle ou leur importance diminuera dans la plupart des pays tant que les systèmes éducatifs pourront absorber des effectifs encore plus élevés. Il faudrait donc s'efforcer de limiter les effets négatifs imputables à l'emploi d'instruments inappropriés, même si les effets négatifs induits par les enjeux de l'examen demeurent. Plusieurs orientations sont proposées dans l'Encadré 6 pour améliorer la qualité des examens nationaux (Kellaghan & Greaney, 1992).

Encadré 6. Orientations pour l'amélioration des examens nationaux

- Les évaluations doivent refléter le programme scolaire et, faute de pouvoir évaluer tout ce programme dans le cadre d'un examen unique, il faut que les domaines évalués soient des domaines considérés comme essentiels. Si l'on veut que le niveau scolaire des élèves corresponde au contenu des épreuves, il est capital d'explicitier clairement les objectifs visés et le contenu des épreuves. Les objectifs doivent refléter la contribution apportée au développement et à l'épanouissement des élèves à long terme en termes de connaissances théoriques et pratiques (Le Mahieu, 1984).
- Dans la mesure du possible, différents modes d'évaluation (écrite, pratique, orale par ex.) seront utilisés pour refléter les objectifs des programmes scolaires.
- Les questions posées dans les épreuves ne doivent pas exclusivement faire appel à la mémoire ou à la reconnaissance, mais évaluer également des qualités supérieures en termes de facultés cognitives plus complexes (par ex. compréhension, synthèse, application).
- Les examens doivent évaluer la capacité des élèves à utiliser leurs connaissances non seulement dans des contextes scolaires, mais aussi extrascolaires.
- Les examens doivent, tant sur le plan du contenu que sur le plan du degré de difficulté, jouer leur fonction d'authentification (délivrance de diplômes, certificats) et permettre une évaluation représentative des compétences des élèves qui n'obtiendront pas une qualification de haut niveau.

...

...

- Les performances aux examens doivent être analysées et les conclusions de cette analyse retransmises aux écoles et autres partenaires (autorités responsables des programmes scolaires, etc.).
- Le nombre d'enseignants participant à la conception des sujets et à la notation des épreuves doit être aussi élevé que possible; en effet, la participation à ces activités leur donnera une meilleure perception des exigences demandées et leur permettra d'adapter leur enseignement en conséquence.
- La question d'une éventuelle prise en compte des évaluations d'enseignants dans les notes attribuées à leurs élèves aux examens nationaux mérite d'être étudiée.
- Le contenu et la forme des questions posées aux examens doivent être exempts de tout préjugé sexiste, ethnique et géographique. Ainsi, la langue employée ne doit revêtir aucun caractère sexiste (comme dans les noms utilisés par exemple), les sujets doivent intéresser aussi bien les filles que les garçons et exprimer des situations susceptibles d'être familières aux élèves, filles et garçons. Il est à noter que les questionnaires à choix multiple tendent, semble-t-il, à favoriser les garçons.

(Kellaghan, & Greaney, 1992)

3. EVALUATIONS NATIONALES

88. Contrairement aux examens nationaux qui sont depuis longtemps une caractéristique inhérente aux systèmes éducatifs, les évaluations nationales (parfois appelées évaluations du système, évaluations d'apprentissage et, de façon moins correcte, évaluations des résultats d'apprentissage) sont relativement nouvelles (Greaney & Kellaghan, 1996a; Kellaghan, 2003; Kellaghan & Greaney, 2001a, b). Par évaluation nationale, on entend un procédé utilisé pour illustrer le niveau, non pas d'élèves individuellement, mais de l'ensemble d'un système éducatif ou d'une partie clairement définie d'un système (les élèves en quatrième année de primaire ou les élèves âgés de 11 ans par exemple). La pièce maîtresse de l'évaluation est la collecte de données dans les écoles. Habituellement les élèves répondent à des instruments et à des questionnaires d'évaluation en groupe. Les enseignants peuvent aussi être invités à remplir des questionnaires dans lesquels ils donnent les informations qu'ils jugent utiles pour l'interprétation des résultats de leurs élèves (Kellaghan & Greaney, 2001). Les principaux éléments d'une évaluation nationale sont répertoriés dans l'*Encadré 7*.

Encadré 7. Principaux éléments d'une évaluation nationale

- Le ministère de l'éducation désigne un organisme d'exécution, soit au sein du ministère, soit indépendant et externe (département universitaire ou organisme de recherche par exemple), et fournit les moyens de financement nécessaires.
- Le ministère, le cas échéant en concertation avec les principaux partenaires de l'éducation (représentants d'enseignants, spécialistes des programmes scolaires, personnalités du monde des affaires, parents par ex.), définit les exigences pratiques auxquelles doit satisfaire l'évaluation.
- Le ministère de l'éducation, ou une commission d'organisation nommée par lui, définit la population à évaluer (élèves de quatrième année par ex.).
- Le domaine de performances à évaluer est défini (par exemple, la lecture, les maths).
- L'organisme d'exécution prépare les tests de niveau, les questionnaires annexes et les manuels administratifs.
- Les épreuves et les documents complémentaires font l'objet d'une expérimentation préliminaire, puis ils sont réexaminés afin d'en contrôler l'adéquation avec les programmes scolaires et la faisabilité technique.
- L'organisme d'exécution sélectionne l'échantillon cible (ou la population d'écoles/d'étudiants), assure l'impression des documents et informe les écoles choisies.
- Les personnes chargées de faire passer les tests (enseignants, inspecteurs scolaires ou étudiants de l'enseignement supérieur) sont formées par l'organisme d'exécution.
- Les instruments de mesure (épreuves et questionnaires) sont administrés dans les écoles à une date définie.
- Les instruments de mesure sont collectés, renvoyés à l'organisme d'exécution, triés et préparés aux fins d'analyse.

(Kellaghan & Greaney, 2001, p. 35).

89. Les évaluations nationales ont été instaurées parce que les données de l'éducation sur les intrants éducatifs que l'on collectait par le passé étaient souvent peu pertinentes ou difficilement utilisables par les planificateurs de l'éducation (Kudjoh &

Mingat, 1993). Les évaluations nationales avaient pour but de résoudre ce problème en apportant des informations sur les « produits » ou les « résultats » de la scolarité (niveau des élèves, inégalités du système par exemple), informations susceptibles d'être exploitées en lien avec les données sur les intrants afin de servir de base au processus d'élaboration de politiques et de prise de décision dans le secteur de l'éducation. Les évaluations nationales entendaient donc doter les responsables des politiques et des décisions d'informations fiables et pertinentes sur l'état du système éducatif, ainsi que sur ses forces et ses faiblesses, et qui puissent se prêter à l'analyse et à l'interprétation. Ces informations devaient a priori être différentes des données émanant d'autres sources. Elles devaient ainsi différer des données issues des examens nationaux qui sont révélatrices des performances scolaires, mais uniquement des candidats aux examens. Elles devaient par ailleurs se différencier des données produites par les recherches et les études sur le secteur éducatif, lesquelles portent généralement sur le court terme; les évaluations nationales en revanche, lorsqu'elles sont intégrées dans la gestion et l'administration globales du système éducatif, fournissent des informations qui sont utilisables aux fins d'une analyse à plus long terme du secteur éducatif.

90. Parmi les évaluations nationales les plus anciennes et les plus connues, on peut citer le cas du Royaume-Uni qui a mis en place différents mécanismes depuis 1948, celui des Etats-Unis d'Amérique où la première évaluation du genre remonte à 1969 et celui de la France qui a instauré un système similaire en 1979. Le Chili a également lancé en 1978 une évaluation nationale sur une longue période. Tous très élaborés, ces systèmes fonctionnent sur la base d'une collecte fréquente de données. Les systèmes d'évaluation mis en place dans les autres pays sont en général plus récents et moins perfectionnés (Kellaghan & Greaney, 2001b).

91. Dans les pays économiquement moins développés, l'instauration des évaluations nationales ne date que des années 90. Dans la plupart des cas, l'élément déclenchant a, semble-t-il, été la Déclaration de la Conférence mondiale sur l'Education pour tous qui soulignait la nécessité d'identifier le lien effectif entre l'apprentissage des élèves et les offres d'éducation (les produits de l'éducation) et de disposer d'informations plus précises afin d'améliorer l'efficacité de la planification et de la gestion de l'éducation. Face au manque de données disponibles en la matière, le besoin d'organiser des évaluations nationales et de mettre en place les capacités nécessaires à leur réalisation s'est fait sentir.

3.1. Informations recherchées dans le cadre des évaluations nationales

92. Toutes les évaluations nationales ont pour but d'apporter des réponses à l'une au moins des questions ci-dessous.

- Quelle est la qualité d'apprentissage des élèves dans le système éducatif (par rapport aux attentes globales, aux objectifs de l'EPT, aux finalités des programmes scolaires ou à la préparation de leur avenir par exemple) ?
- Peut-on discerner des points forts ou des points faibles particuliers dans les connaissances théoriques et pratiques acquises par les élèves ?
- Existe-t-il des écarts entre les résultats scolaires de différents sous-groupes d'une population ? Y a-t-il par exemple des disparités de niveau entre garçons et filles, entre élèves de région rurale et élèves de région urbaine, entre élèves de langue différente, entre élèves appartenant à des groupes ethniques différents, entre élèves issus de

régions différentes du pays ou entre élèves qui abandonnent prématurément leur scolarité ou qui redoublent des classes ?

- Y a-t-il un lien, et lequel, entre les résultats scolaires et les caractéristiques du contexte d'apprentissage (ressources des écoles, préparation et qualifications des enseignants, type d'établissement scolaire par exemple) ou l'environnement familial ou communautaire des élèves ?
- Les performances des élèves varient-elles avec le temps? Cet aspect peut être particulièrement important si une réforme majeure du système est engagée (augmentation des effectifs, introduction de nouvelles matières, instauration de nouveaux programmes par exemple). (Kellaghan, 2003 ; Kellaghan & Greaney, 2001b)

3.1.1. Divers usages des informations fournies par une évaluation nationale

93. Les données fournies par une évaluation nationale ont été utilisées à différentes fins, notamment pour :

- attirer l'attention des responsables politiques, des décideurs et de l'opinion publique sur la nécessité d'améliorer l'efficacité de l'éducation (voire d'augmenter le volume de ressources) si l'on souhaite promouvoir le développement social et économique ;
- justifier le redéploiement des ressources discrétionnaires; il est à noter cependant que la souplesse d'affectation des ressources peut, dans certains cas, être réduite (par ex. lorsque l'essentiel des ressources sert à payer les salaires des enseignants) ;
- apporter une aide aux décisions à prendre par les pouvoirs publics (concernant la politique de redoublement par ex.) ;
- éviter les « mauvaises » décisions des pouvoirs publics (retrait du financement de la formation continue par ex.) qui ne prennent souvent en compte qu'un seul aspect de la politique, en analysant les problèmes dans une perspective plus large ;
- proposer une affectation de ressources plus rationnelle, inspirée par les conclusions d'analyses relatives au rapport entre d'autres intrants et le niveau scolaire des élèves ;
- introduire une gestion plus efficace par le biais d'une plus grande transparence (voire, dans certains cas, de la concurrence) ;
- utiliser les effets secondaires de l'évaluation (comme il a été proposé dans le cas des examens nationaux) pour veiller à ce que les enseignants enseignent certaines matières. Cela n'est possible que si toutes les écoles prennent part à l'évaluation et que des sanctions sont liées aux performances (par ex. au Chili; voir Himmel, 1996; Schiefelbein, 1993).

3.2. Différences entre les évaluations nationales

94. Si les évaluations nationales ont pour but commun de décrire le niveau scolaire des élèves dans le système éducatif, elles mettent en œuvre des procédures très différentes selon le cas (voir *Encadré 8*). Ces différences de stratégie influent sur le mode d'exploitation des données qu'elles fournissent. A titre d'exemple, il n'est pas logique de cibler une évaluation sur les performances des écoles (en publiant les résultats par ex.) si celle-ci ne porte que sur un échantillon d'écoles. De même, utiliser les résultats d'une

évaluation pour apprécier les performances des élèves ou des enseignants n'a aucun sens si chaque élève ne passe qu'une partie des épreuves (une sur 9 dans certaines évaluations).

3.3. Initiatives d'évaluations nationales en Afrique

95. On compte quatre grandes catégories d'évaluation nationale en Afrique. Trois d'entre elles concernent des initiatives similaires menées dans plusieurs pays : le projet "Monitoring Learning Achievement" (MLA) ; le projet du Consortium de l'Afrique australe pour le pilotage de la qualité de l'éducation (SACMEQ) ; le Programme d'Analyse des Systèmes Educatifs des Pays de la CONFEMEN (PASEC). La quatrième regroupe des évaluations nationales indépendantes réalisées dans quelques pays.

Encadré 8. Différences relatives aux procédures d'évaluations nationales

- Les évaluations portent, dans la majorité des cas, sur un échantillon d'écoles/d'élèves; dans certains cas, la totalité ou presque des élèves appartenant à une tranche d'âge ou à un niveau de classe donnée est évaluée.
- Un élève a la possibilité de ne passer qu'une partie des nombreuses épreuves prévues dans le cadre de l'évaluation, ce qui permet de couvrir un large éventail du programme scolaire sans que les élèves perdent du temps à passer toutes les épreuves (Brésil, Etats-Unis d'Amérique, Irlande par ex.); ou encore tous les élèves peuvent passer les mêmes épreuves (Ouganda, Zambie par ex.).
- Le but d'une évaluation peut, le cas échéant, être de fournir des indications sur les écoles (voire sur les enseignants ou les élèves). Le fait que l'on dispose d'informations sur chaque école peut conférer à une évaluation un enjeu aussi important qu'un examen public (Angleterre par ex.).
- Les évaluations diffèrent au niveau du mode de présentation, plus ou moins complexe, des performances des élèves. Les résultats peuvent être présentés sous la forme d'un simple pourcentage de bons résultats; les notes des élèves peuvent aussi être étalonnées d'après un système de modélisation des réponses aux questions.
- Les évaluations diffèrent au niveau du degré de détails utilisé pour décrire les performances (par ex. par sous-domaines d'un domaine de mathématiques ou de calcul plus large) et au niveau du lien éventuellement recherché (lien de causalité dans certains cas) entre les résultats des élèves et différents paramètres (type d'école, caractéristiques de la communauté, par ex.).
- Les évaluations diffèrent au niveau de la façon dont les données ainsi obtenues sont intégrées dans d'autres aspects du système éducatif.

96. Lancé en 1992, le projet MLA est une initiative conjointe de l'UNESCO et de l'UNICEF. Il fait partie intégrante de l'évaluation EPT qui a pour but de dresser un tableau d'ensemble des progrès accomplis par les pays sur la voie des objectifs de l'EPT (Chinapah, 1997). Il répond en particulier au besoin d'étudier sur une longue période la manière dont les étudiants acquièrent des connaissances théoriques et pratiques, une capacité de raisonnement et des valeurs (Article 4 de la *Déclaration mondiale sur l'éducation pour tous*, 1990).

97. Pour la phase MLA I, l'évaluation a porté sur les performances des élèves de quatrième année de primaire en lecture et écriture, calcul et compétences liées à la vie quotidienne (intérêt et connaissances en termes de santé, nutrition, hygiène). Pour la phase MLA II, l'évaluation était ciblée sur les performances d'élèves de huitième année en mathématiques et sciences; des épreuves communes ont eu lieu dans tous les pays. Ont

également été collectées des données concernant les antécédents scolaires des élèves, les caractéristiques des écoles et les antécédents familiaux.

98. Pour l'heure, des évaluations MLA ont été pratiquées dans plus de 70 pays, dont 47 en Afrique (UNESCO, 2003a). Quarante pays africains ont pris part au projet MLA I et 11 au projet MLA II. Au Nigeria, les résultats ont été classés par Etat. Une approche nationale a été adoptée pour tirer parti des compétences locales et développer les capacités des pays participants. En mars 2003, les rapports sur les évaluations MLA I de 18 pays subsahariens et sur les évaluations MLA II de deux pays avaient été publiés.

99. Outre les rapports nationaux, un rapport distinct sur l'évaluation MLA I de 11 pays a été rédigé (Botswana, Ile Maurice, Madagascar, Malawi, Mali, Maroc, Niger, Ouganda, Sénégal, Tunisie, Zambie) (Chinapah et al, 2000). Il en ressort que seuls quatre de ces pays ont atteint l'objectif d'apprentissage fixé à Jomtien (c'est-à-dire que 80 % des apprenants doivent posséder un certain niveau de compétences d'apprentissage) pour les élèves de quatrième année dans le domaine des compétences essentielles à la vie quotidienne; deux pays pour ce qui est de la lecture/l'écriture; et aucun pays pour ce qui est du calcul. Les écarts entre garçons et filles sont faibles dans tous les pays. A l'exception de l'île Maurice, les élèves du secteur privé ont de meilleurs résultats que ceux du secteur public dans les trois domaines d'apprentissage évalués. Dans la plupart des pays, on observe un lien entre la capacité des parents à aider les apprenants à faire leurs devoirs et les résultats des élèves.

100. Ile Maurice Le SACMEQ est un groupement volontaire qui rassemble 15 ministères de l'éducation d'Afrique australe et orientale et qui travaille en étroite collaboration avec l'Institut international de planification de l'éducation (IIPE) à Paris aux fins de créer, par le biais d'une formation conjointe, les capacités institutionnelles nécessaires pour mener des recherches communes sur la politique éducative (Ross et al, 2000; UNESCO, 2003b). Entre 1995 et 1998, huit ministères de l'éducation ont collecté des données dans le cadre de la phase SACMEQ I sur des indicateurs de référence relatifs aux intrants éducatifs, aux conditions générales de scolarité, à une répartition équitable des ressources humaines et matérielles et aux niveaux d'alphabétisation des élèves de sixième année. L'évaluation a porté sur des enseignants et sur des élèves, sauf à l'île Maurice et en Afrique du Sud. Quinze pays ont pris part à la phase SACMEQ II entre 1999 et 2002 : Afrique du Sud, Botswana, Kenya, Ile Maurice, Lesotho, Malawi, Mozambique, Namibie, Ouganda, Seychelles, Swaziland, Tanzanie (continent), Tanzanie (Zanzibar), Zambie, Zimbabwe. Sept rapports nationaux ont été publiés et 14 sont en préparation.

101. L'une des missions principales du SACMEQ est de promouvoir le renforcement des capacités en dotant les planificateurs de l'éducation des pays membres des compétences techniques requises pour piloter et évaluer les conditions de scolarisation et la qualité de l'éducation. Grâce aux systèmes d'informations efficaces et accessibles mis à la disposition des décideurs, il vise également à promouvoir la participation des partenaires et à accroître la transparence. Cette approche a notamment la particularité d'offrir aux planificateurs de l'éducation une formation de type apprentissage par l'action en les impliquant dans la conduite des études.

102. Comme pour le MLA, les résultats des évaluations SACMEQ indiquent que les systèmes éducatifs ne satisfont pas aux critères de performances des ministères. A titre d'exemple, en Namibie et en Zambie, moins de 30 % des élèves de sixième année ont atteint le niveau minimum spécifié en lecture/écriture (UNESCO, 2000a). Même si, au Zimbabwe, plus de 50 % possèdent le niveau « minimum », ce chiffre n'a pas évolué au fil des années (Machingaidze, Pfukani, & Shumba, n.d.). Dans l'ensemble, on constate

des écarts de niveau non négligeables d'une région à l'autre et d'un type d'école à l'autre dans un même pays. Les inégalités entre garçons et filles sont mineures (Mioko, 1998).

103. Créé en 1991 par les pays francophones pour répondre à la Déclaration de Jomtien, le PASEC est un réseau destiné à l'échange d'informations sur des instruments et des résultats d'évaluation de l'éducation utilisés dans les pays francophones en Afrique. Il encourage les hauts responsables des décisions et d'autres partenaires à participer aux débats sur les questions de politique éducative et souligne la nécessité pour les pouvoirs publics de s'appuyer sur des données fiables afin de prendre des décisions et d'élaborer des programmes d'action réalistes (notamment en termes de prévision des délais et des coûts) (Kulpoo & Coustère, 1999). Au départ, les élèves de la deuxième et de la sixième année de l'école primaire ont été évalués en français et en mathématiques. Maintenant, les élèves de la deuxième à la sixième année sont évalués en français et en mathématiques. Depuis 1994, les mêmes épreuves sont utilisées dans tous les pays. L'évaluation s'étend en outre aux élèves et aux enseignants d'écoles et d'origines différentes. Douze pays ont pris part au PASEC : le Burkina Faso, le Cameroun, la République centrafricaine, le Congo, la Côte d'Ivoire, Djibouti, la Guinée, Madagascar, le Mali, le Niger, le Sénégal et le Togo. On estime à environ 50 000 dollars américains le coût d'une évaluation PASEC ou SACMEQ (Kulpoo et Coustère, 1999).

104. L'une des différences entre le PASEC et d'autres évaluations nationales pratiquées en Afrique est que, dans certains pays, les élèves sont évalués vers le début (novembre) et vers la fin (mai) de l'année scolaire. Cette formule permet donc d'avoir un aperçu de l'évolution ou de la « valeur ajoutée » pendant l'année, même si elle ne prend en compte que les élèves qui ont en quelque sorte survécu au système. Elle met en évidence la nette orientation du programme vers la recherche, caractéristique déjà très perceptible lors de sa conception, et elle s'efforce d'identifier les « liens de causalité » susceptibles de servir de base à l'élaboration de politiques destinées à améliorer l'efficacité et le rendement de l'éducation. Diverses tentatives ont été faites pour déterminer « l'impact » des facteurs scolaires intrinsèques (formation des maîtres, effectifs des classes, accessibilité des manuels scolaires) et des facteurs extrinsèques (éducation des parents, distance domicile-école, langue maternelle).

105. Hormis ces trois programmes conçus pour soutenir des évaluations nationales, plusieurs pays ont mené à bien des évaluations nationales axées vers des aspects techniques spécifiques (Burundi, Erythrée, Mali, Namibie, Ouganda, Sénégal, Zambie). En Namibie, une Evaluation nationale de référence des apprenants a été effectuée pour mesurer le niveau d'anglais et de mathématiques d'élèves de quatrième et de septième années (Ministère de l'éducation et de la culture de Namibie/Université d'Etat de Floride/Université de Harvard, 1994). Le niveau de lecture et de compréhension orale a été évalué à Oshindonga dans un sous-ensemble d'écoles de la région nord. Le but était d'apporter une aide aux responsables des politiques pour l'affectation de ressources à des écoles peu performantes. Il est apparu que le niveau d'anglais escompté était trop élevé et que les documents relatifs aux programmes scolaires devaient être remaniés.

106. En Erythrée, le ministère de l'éducation générale (1999) a, dans le cadre de l'évaluation EPT 2000, organisé une évaluation dans six régions en vue de déterminer si les élèves possédaient les compétences de base en langue maternelle (première année) et en anglais et mathématiques (quatrième année) spécifiées dans les programmes officiels. L'évaluation a permis de mettre en lumière quelques problèmes spécifiques relatifs aux programmes scolaires (valeur des places, difficultés de vocabulaire en mathématiques par ex.), de constater que les garçons avaient globalement de meilleurs résultats que les filles et de dégager des conclusions en matière de formation des enseignants et de méthodes pédagogiques.

107. Ces évaluations nationales réalisées en Afrique présentent des points communs. Ce sont des évaluations auxquelles sont associés des planificateurs et des gestionnaires et qui ont pour but de fournir des informations utiles à l'élaboration de politiques. Ce n'était pas le cas du PASEC qui, à l'origine, était essentiellement axé vers la recherche et les études techniques. Mais, avec le temps, le projet a suscité davantage d'intérêt et de mobilisation parmi les responsables des ministères. Dans toutes ces évaluations, le renforcement des capacités et du couple décideurs/chercheurs compte parmi leurs objectifs prioritaires. Elles utilisent la technique d'échantillonnage. Elles s'appuient sur un modèle du système éducatif de type entrées-sorties et visent à dégager les paramètres qui interviennent dans les résultats scolaires.

3.4. Impact des évaluations nationales

108. L'impact produit dans divers domaines par tous ces programmes d'évaluation a fait l'objet de rapports émanant d'équipes d'évaluation ou d'organes connexes. Les résultats du SACMEQ ont figuré au centre de différents débats sur les politiques et leur formulation, notamment dans le cadre de commissions présidentielles et nationales au Zimbabwe et en Namibie; de réunions ministérielles sur la politique éducative à Zanzibar; d'études nationales sur le secteur éducatif en Zambie; et d'études de plans directeurs nationaux de l'éducation à l'île Maurice (Ross et al, 2000). Les résultats du projet MLA ont été exploités pour préparer les réformes de l'éducation au Malawi et en Ouganda, quant aux résultats du PASEC, ils ont été évoqués dans des rapports d'examen par pays au Burkina Faso et au Niger.

109. En outre, les données fournies par ces programmes ont, semble-t-il, contribué au renforcement des capacités. A la suite du projet MLA, plusieurs pays ont effectué leur propre évaluation (Ile Maurice, Madagascar, Mali, Maroc). Au Kenya et au Zimbabwe, les méthodes de tri de données utilisées par le SACMEQ ont été adaptées aux données du recensement des effectifs scolaires. Dans les pays participant au PASEC, on a enregistré une amélioration des capacités à construire des épreuves, ainsi qu'à concevoir et réaliser des enquêtes.

110. Les résultats des évaluations nationales concernant les redoublements, les effectifs et la disponibilité des manuels scolaires ont été utilisés pour aider à l'élaboration des politiques dans des pays participant au PASEC. La question de savoir s'il était justifié de procéder ainsi reste toutefois sans réponse. L'importance donnée à la disponibilité des manuels scolaires est moindre dans les analyses ultérieures des données relatives au Sénégal que dans les premières analyses (Naumann & Wolf, 2001).

111. Les résultats des évaluations nationales ont également suscité des débats à l'échelon national. Ainsi, à l'île Maurice, les données SACMEQ ont été utilisées lors d'un débat national sur les stratégies de nature à limiter l'effet négatif du Certificat d'études primaires et à améliorer le rapport de l'élève avec le livre. Aux Seychelles, les données SACMEQ ont donné matière à discussion sur la répartition par niveaux d'aptitudes.

112. Dans plusieurs pays, les résultats SACMEQ ont conduit les administrateurs de l'éducation à réviser le volume de ressources affectées à l'éducation. Au Kenya, on a instauré des indices de référence concernant l'octroi des ressources (nombre de bureaux par élève, nombre de livres par élève). Au Zimbabwe, on a réparé des bâtiments scolaires, accordé des fonds spéciaux pour les fournitures scolaires et institué des programmes de formation à la gestion des bibliothèques scolaires.

113. Dans certains systèmes éducatifs, grâce à l'expérience acquise avec le MLA, l'attention portée dans les programmes scolaires aux compétences indispensables à la vie courante a augmenté. Néanmoins, l'impact sur les pratiques scolaires reste à démontrer.

114. Les résultats des évaluations ont par ailleurs été utilisés pour justifier auprès de bailleurs de fonds des demandes d'aide substantielle en vue de pallier les faiblesses constatées dans le système. En Guinée, par exemple, les résultats de l'évaluation PASEC ont conduit le gouvernement et la Banque mondiale à élaborer un programme visant à promouvoir l'enseignement en langue locale dans les petites classes, avant de procéder à l'instauration progressive du français.

115. Il convient toutefois de ne pas surestimer les effets des évaluations nationales. A l'île Maurice, selon une équipe d'analyse externe, les mesures pratiques suggérées dans trois évaluations nationales distinctes n'ont pratiquement eu aucune suite. Interrogés sur l'utilisation des résultats des évaluations nationales pour les besoins de la présente étude, de hauts responsables de six pays (Afrique du Sud, Ethiopie, Malawi, Niger, Nigeria et Ouganda) ont déclaré que les résultats des évaluations avaient été occasionnellement mentionnés dans les media. Par contre, ces résultats n'avaient été évoqués dans les débats parlementaires d'aucun pays. Les résultats n'avaient été utilisés que dans un seul pays pour justifier l'octroi de ressources supplémentaires aux écoles. Dans quatre pays, les résultats ont suscité des échanges entre les autorités responsables des programmes scolaires, mais le retour d'informations vers les enseignants/écoles n'a eu lieu que dans deux pays seulement et le retour d'informations vers les éditeurs de manuels scolaires dans un seul pays. Dans deux pays sur cinq (Ethiopie, Nigeria), les personnes interrogées ont déclaré que les résultats des évaluations nationales n'avaient pas été utilisés pour la formulation de la politique éducative.

3.5. Questions relatives aux évaluations nationales

116. La décision de réaliser une évaluation nationale soulève plusieurs questions. Certaines doivent être résolues avant le début de l'évaluation; d'autres concernent l'utilisation et l'interprétation des données. Le présent chapitre aborde ces différents aspects en y apportant des éléments d'informations complémentaires issus de l'expérience internationale.

- *Domaine de compétences.* Un choix s'impose quant au domaine de compétences à évaluer. Les domaines les plus couramment analysés dans les évaluations pratiquées de par le monde sont la langue d'enseignement des élèves et les mathématiques, plus rarement les matières scientifiques et les sciences humaines. Une autre question qui se pose est de savoir si les compétences évaluées se rapportent au programme scolaire imposé ou à une évaluation des connaissances théoriques et pratiques dont les élèves auront besoin au terme de leur scolarité.
- *Population.* Certaines évaluations nationales s'adressent à des élèves appartenant à un ou deux niveaux du primaire, tandis que d'autres concernent les élèves de tous les niveaux. Dans le secondaire, si les évaluations sont moins fréquentes, elles existent néanmoins dans plusieurs pays.
- *Tous les élèves ou un échantillon d'élèves.* Dans certains systèmes d'évaluation nationale (au Chili par ex.), tous les élèves (ou presque) des classes concernées sont évalués. Les informations ainsi obtenues couvrent donc la totalité des écoles et des enseignants, de même que la quasi-totalité des élèves. Cette procédure est appliquée lorsqu'il s'agit de préparer des actions à l'échelon de l'école (ou des enseignants), comme la publication des résultats ou l'affectation de soutien et de ressources à des écoles peu performantes. L'évaluation d'un échantillon d'élèves est plus courante. Son coût est par ailleurs nettement inférieur à celui d'une évaluation générale de tous les élèves d'un système. Par contre, les informations obtenues sont utilisées plus pour l'élaboration de politiques que pour des actions ciblées vers des écoles déterminées.

- *Comparaisons entre régions ou districts.* De nombreux pays sont intéressés par les analyses comparatives des performances des élèves d'une région à l'autre ou d'un district à l'autre. Mais bien souvent, les échantillons employés sont trop réduits pour permettre des comparaisons valables.
- *Une même épreuve pour tous les élèves ou une épreuve pour chaque élève portant sur une partie seulement des questions.* Dans plusieurs évaluations nationales et internationales, les élèves ne répondent qu'à une partie seulement des questions de l'épreuve; ceci permet de couvrir une plus grande partie du programme scolaire ou de laisser davantage de place à l'évaluation de la compréhension écrite sans augmenter la charge de travail des élèves. Dans d'autres évaluations, tous les élèves répondent aux mêmes questions. Si la formule dans laquelle les élèves ne répondent qu'à une partie des questions offre des avantages, elle comporte aussi des inconvénients, notamment pour les pays qui lancent un programme d'évaluation nationale. Outre sa complexité, la notation et le barème de notations sont également plus compliqués, sans parler de la difficulté pour analyser les résultats relatifs aux élèves ou aux écoles.
- *Niveaux.* Etablir des « niveaux de compétences », des « niveaux d'aptitudes » ou des « niveaux de maîtrise » ("minimum » ou « souhaitable »), même si les instruments d'évaluation sont définis sur la base de critères de référence, est beaucoup plus problématique qu'il n'y paraît. Il n'existe aucune règle simple qui permette de savoir le seuil à partir duquel un niveau peut être considéré comme atteint ou d'affirmer que les élèves ayant des notes supérieures à tel seuil se situent à un niveau de performances nettement différent des élèves ayant des notes juste inférieures à ce seuil. A l'heure actuelle, la fixation de niveaux repose en général sur un jugement des individus. Cependant, comme les procédures appliquées sont variables et que les jugements sont toujours quelque peu arbitraires (voir Cizek, 2001), la méthode employée pour fixer des niveaux doit dans tous les cas être explicitée. En outre, il convient d'examiner le lien entre les critères de compétences appliqués dans différents systèmes d'évaluation. A chaque procédure de fixation de niveaux correspond une valeur de seuil différente même si chacune repose sur une base nationale. Les résultats des examens nationaux et des évaluations nationales peuvent mettre en évidence de fortes disparités qu'il importe de combler. Au Sénégal, l'incohérence entre les notes des élèves aux épreuves du PASEC et les pratiques des enseignants concernant le passage dans la classe supérieure a été un sujet de préoccupations. Des élèves qui avaient relativement bien réussi les épreuves PASEC n'étaient pas autorisés à passer dans la classe supérieure, alors que d'autres qui avaient eu des résultats médiocres passaient dans la classe supérieure (B. Moll, communication personnelle, mai 2003). Des écarts peuvent également apparaître entre les appréciations faites sur les performances aux examens nationaux (Kellaghan & Farrell, 1998) et les appréciations faites sur les performances aux évaluations nationales. Au Lesotho, où quatre élèves sur cinq ont réussi l'examen du certificat d'études primaires, moins d'un sur six a atteint le niveau minimum requis à l'évaluation nationale de lecture/écriture. Au Malawi, près de quatre élèves sur cinq ont été admis à l'examen du certificat d'études primaires, mais un sur cinq seulement est considéré dans une évaluation nationale comme possédant le niveau minimum de maîtrise requis. En Ouganda, environ 70 % des élèves ont subi avec succès l'examen du certificat, mais un tiers seulement a obtenu le niveau de maîtrise minimum requis dans une évaluation nationale. Il est vrai que les chiffres relatifs à l'examen et à l'évaluation nationale ne concernent pas les mêmes groupes d'élèves. Toutefois, les disparités sont trop grandes pour n'être pas représentatives de réels écarts entre les niveaux requis dans les examens nationaux et dans les évaluations nationales.
- *Suivi dans le temps.* Si l'on veut suivre l'évolution des niveaux scolaires dans la durée, plusieurs conditions sont à remplir : emploi des mêmes instruments d'évaluation (ou, en cas de changement, il faut s'assurer que les nouveaux instruments employés sont

équivalents); utilisation d'échantillons aléatoires de taille appropriée (cas d'une évaluation qui ne porte pas sur la totalité d'une population) ; et transparence dans la définition des « niveaux ». Il convient de remarquer qu'il n'est pas toujours possible de garantir l'équivalence des instruments (cas d'un nouveau programme scolaire qui comporte des éléments plus pertinents par rapport à la vie des élèves ou qui favorise la compréhension, l'analyse et l'application).

- *Interprétation des liens de corrélation.* Divers rapports relatifs à des évaluations nationales ont tendance à dégager des liens de causalité entre des intrants du système éducatif (caractéristiques de l'école ou de l'élève par ex.) et les performances des élèves. Or, il est rare de pouvoir établir des liens de causalité valables à partir d'enquêtes transversales. Qui plus est, les estimations sont parfois peu fiables du fait que le nombre d'écoles/d'élèves dans certaines catégories (dans des districts, dans des régions rurales par ex.) est faible et que les méthodes d'analyse sont inadaptées.
- *Renforcement des capacités.* D'après les données d'expérience disponibles à ce jour, il apparaît que les personnes compétentes et dûment formées en matière d'évaluation nationale restent rarement longtemps en fonction. Elles sont en général promues à des postes plus élevés.
- *Participation de partenaires.* Dans certains systèmes éducatifs, il est parfois difficile de garantir la représentativité de la communauté élargie des partenaires (formateurs d'enseignants, syndicats d'enseignants, groupes minoritaires par ex.) dans la coordination et la réalisation d'une évaluation nationale. L'opposition d'un groupe de partenaires peut être nuisible à une évaluation.

3.5.1. Conditions et pratiques requises pour développer et institutionnaliser les capacités d'exécution d'une évaluation nationale

117. Différentes conditions et pratiques doivent être respectées pour mener à bien et institutionnaliser une évaluation nationale (Chapman & Mählck, 1993 ; Kellaghan & Greaney, 2001 ; Marope, 1999 ; Samoff, 1999).

- Le besoin de disposer d'informations (y compris sur les performances des élèves) et le besoin de développer les capacités requises pour les obtenir doivent être pris en compte. Ceux qui ont la responsabilité de prendre des décisions concernant une évaluation nationale et de fournir les ressources nécessaires pour la mettre en œuvre (ministre ou secrétaire général par ex.) doivent être conscients de l'importance de disposer d'informations pour être à même de repérer les problèmes et de connaître les politiques et les pratiques. A défaut de connaître l'existence d'un manque de capacités et d'être prêts à le combler, leurs efforts pour renforcer les capacités d'exécution d'une évaluation nationale et ce, dans une perspective durable, risquent d'être voués à l'échec.
- L'accès aux connaissances théoriques et pratiques nécessaires pour concevoir, organiser et analyser une évaluation nationale doit être garanti. A ce titre, une formation (et une aide financière correspondante) est à prévoir.
- Le soutien de partenaires, en particulier des responsables des décisions, est indispensable. Parmi les signes d'un manque de mobilisation institutionnelle, on peut citer l'affectation à un autre poste d'une personne ayant suivi une formation sur l'organisation des évaluations nationales ou bien l'octroi, aux personnes chargées d'une évaluation, du temps ou des ressources insuffisants pour accomplir leur tâche.
- Une évaluation nationale doit être intégrée dans le contexte de la politique éducative, dans les structures existantes, dans les processus d'élaboration des politiques et des décisions et dans les circuits d'affectation des ressources. Elle ne doit pas être perçue comme un projet ou une opération distincte qui prendra fin au terme de l'évaluation. On peut espérer qu'en procédant ainsi, les résultats de l'évaluation nationale pourront remonter jusqu'aux responsables de la formulation et de l'application des politiques et ne subiront pas le sort de certaines études sectorielles dont les conclusions ne sont jamais parvenues entre les mains des responsables des pouvoirs publics.
- Tous les efforts doivent être faits pour que les évaluations nationales soient alignées sur d'autres mécanismes d'orientation pédagogique existants dans le système éducatif, non seulement avec d'autres systèmes d'évaluation (examens nationaux, évaluations à l'échelon de la classe), mais aussi avec les programmes scolaires, la formation des maîtres, le renforcement des capacités des écoles et les mesures destinées à aplanir les injustices.
- Les Systèmes de gestion de l'information en éducation (EMIS, « Educational Management Information Systems »), lorsqu'ils existent et qu'ils fonctionnent, doivent être adaptés de façon à incorporer des données relatives à la qualité des performances des élèves et aux intrants. Le but en l'occurrence est d'institutionnaliser les données d'évaluation et, par là même, d'en faciliter l'accès aux décideurs et aux administrateurs.
- Les autorités éducatives doivent élaborer des systèmes et des stratégies permettant de communiquer les résultats de l'évaluation nationale à des individus extérieurs au

ministère, en particulier aux enseignants, susceptibles d'être associés à l'application de politiques établies sur cette base.

- Enfin, des propositions concernant l'utilisation des résultats de l'évaluation nationale doivent être faites dans l'optique d'une perception réaliste du rôle des données quantitatives dans la planification et la gestion, dans la mesure où la prise en compte de différents facteurs sociaux, économiques et politiques dans la planification et la distribution des ressources est indispensable.

4. EVALUATIONS INTERNATIONALES

118. Les évaluations internationales présentent de nombreux points communs avec les évaluations nationales pour ce qui est des procédures qu'elles emploient. La principale différence réside naturellement dans le fait qu'elles sont conçues pour être administrées dans plusieurs pays (Beaton et al, 1999; Greaney & Kellaghan, 1996a; Kellaghan & Greaney, 2001).

119. Comme dans le cas des évaluations nationales, des instruments sont mis au point pour apprécier les connaissances théoriques et pratiques des élèves. Cependant, au lieu de refléter le programme scolaire d'un seul et unique système éducatif, ces instruments d'évaluation internationale doivent être utilisables dans tous les systèmes participant à l'évaluation. L'âge ou la classe auquel sont appliqués ces instruments, de même que les procédures de sélection des écoles/élèves, doivent être définis de concert. Pour toutes les études internationales, des échantillons d'élèves ont été utilisés.

120. Depuis les années 60, plus de 60 pays ont participé à des études internationales sur les niveaux de lecture, de mathématiques, de sciences, d'écriture, de littérature, de langues étrangères, d'éducation civique et d'initiation à l'informatique. Ces études donnent des indications sur le niveau scolaire des élèves d'un pays par rapport à celui des élèves d'un autre pays. Elles illustrent également les différentes manières d'aborder les mêmes disciplines dans les programmes scolaires de différents pays. A la lumière de ces informations, un pays peut être conduit à revoir sa politique en matière de programmes scolaires. En effet, le principal impact des études internationales semble être l'impact sur les programmes scolaires (voir par ex. Burnstein, 1993 ; Kellaghan, 1996b; Romberg, 1999).

4.1. Evaluations internationales en Afrique

121. Un petit nombre seulement de pays africains a participé à des évaluations internationales de grande ampleur. Le Ghana, le Nigeria et le Zimbabwe ont pris part à l'étude scientifique de l'AISE en 1983-1984 (Postlethwaite & Wiley, 1992). Aucun pays africain n'a participé à l'étude de l'AISE sur le niveau de lecture et d'écriture; un pays (Afrique du Sud) a participé à la "Third International Mathematics and Science Study" (TIMSS), trois pays (Afrique du Sud, Maroc, Tunisie) à la TIMSS-R et un pays (Maroc) à la "Progress in International Literacy Study" (PIRLS). Il n'y a là rien de surprenant dans la mesure où ces études étaient conçues pour des pays industrialisés. La participation de pays africains à des études internationales qui sont adaptées aux conditions et normes du monde industrialisé n'aurait qu'une valeur limitée.

122. En Afrique du Sud, les élèves ont obtenu des résultats médiocres aux épreuves TIMSS et TIMSS-R. S'agissant du TIMSS-R, la moyenne en mathématiques obtenue par les élèves de huitième année était inférieure à la moyenne obtenue par les élèves les moins performants dans la quasi-totalité des pays participants. Les résultats de l'Afrique du Sud ont permis de comparer les performances par province (Howie, 2001) et d'établir des liens entre les niveaux scolaires et les variables liées aux antécédents des élèves, parmi lesquels la langue maternelle, la situation socio-économique, la race, la langue d'enseignement et l'attitude des élèves (Howie, 2002).

123. Bien que les évaluations MLA, PASEC et SACMEQ n'aient qu'une portée nationale, les rapports qui ont été faits de leurs résultats permettent des comparaisons à l'échelon international. Toutefois, pour que ces comparaisons soient valables, les instruments, les populations cibles, l'échantillonnage et les analyses doivent a priori être

identiques dans tous les pays. En l'occurrence, rien ne permet d'affirmer que ces conditions soient respectées. A titre d'exemple, le projet MLA prévoyait l'élaboration « d'instruments spécifiques par pays » (Chinapah et al, 2000, p. 4). Si les instruments varient d'un pays à l'autre, l'interprétation des notes moyennes pour 11 pays (présentée dans Chinapah et al, 2000, p. 82) est problématique.

4.2. Difficultés liées aux évaluations internationales

124. Plusieurs difficultés ont été relevées en lien avec les études internationales (Kellaghan & Greaney, 2001b). Ce constat est valable pour les études réalisées en Afrique comme ailleurs, même s'il faut reconnaître que les pays en développement sont potentiellement plus exposés à ces difficultés que les pays industrialisés. La première difficulté que l'on peut rencontrer tient au fait que concevoir une procédure d'évaluation permettant de mesurer efficacement les résultats d'un large éventail de programmes scolaires n'est pas chose aisée. S'il existe des points communs entre les divers programmes scolaires appliqués dans le monde, notamment à l'école primaire, on observe des écarts considérables entre les pays industrialisés et les pays en développement quant au contenu de l'enseignement et aux niveaux scolaires des élèves. Plus l'écart entre les programmes scolaires et les niveaux des élèves des pays participant à une évaluation internationale est grand, plus il est difficile et, à dire vrai sans doute impossible, d'élaborer une procédure d'évaluation susceptible de convenir à tous les pays.

125. Une deuxième difficulté peut se présenter lorsque les instruments doivent être traduits dans une ou plusieurs langues. Cette difficulté peut aussi se rencontrer à l'intérieur d'un pays. Dans l'un et l'autre cas, si l'on est amené à comparer les performances évaluées dans différentes langues, il faut savoir que les écarts éventuellement constatés peuvent être imputables à des différences d'ordre linguistique dans le degré de difficulté des épreuves d'évaluation. Ainsi, il est très difficile de garantir que la formulation des questions et l'adéquation culturelle de leur contenu sont équivalentes dans chacune des versions linguistiques d'une épreuve d'évaluation.

126. Une troisième difficulté liée aux évaluations internationales concerne l'équivalence des populations et des échantillons d'élèves évalués entre les différents pays. Cette difficulté apparaît notamment lorsque les taux de redoublement varient d'un pays à l'autre. Là encore, elle est plus fréquente dans les études auxquelles participent des pays industrialisés et des pays en développement. Mais il peut arriver aussi que les écarts liés aux populations et à l'échantillonnage soient plus subtils : ce peut être le cas par exemple lorsque des élèves qui suivent des programmes éducatifs spécialisés ou qui présentent des difficultés d'apprentissage sont exclus dans un pays, mais non dans un autre.

127. Une quatrième difficulté peut provenir du fait que, lors de la notification des résultats d'une évaluation internationale, l'accent est placé sur le classement des pays en termes de notes moyennes obtenues par les élèves, un aspect qui, en général, intéresse tout particulièrement les médias. Or, ces classements peuvent être trompeurs si l'on ne tient pas compte de la signification statistique des écarts moyens entre les performances. Par ailleurs, les classements eux-mêmes ne donnent aucune information sur les multiples facteurs susceptibles d'être à l'origine des écarts de performances d'un pays à l'autre (voir Mislevy, 1995).

128. Enfin, si l'on peut admettre que l'analyse des corrélations entre les intrants scolaires et le niveau des élèves dans certains pays peut s'appliquer à d'autres pays, on ne saurait pour autant en conclure que les pratiques associées à un haut niveau scolaire dans un pays produiront automatiquement les mêmes effets dans un autre. Les corrélations

entre intrants, processus et résultats doivent être analysées dans le contexte de chaque pays (Chapman & Mählck, 1993).

Encadré 9. Evaluations internationales : le cas de l'Afrique du Sud

Le cas de l'Afrique du Sud en matière d'études TIMSS et TIMSS-R illustre les difficultés rencontrées pour mettre en œuvre des évaluations internationales. Howie (1999) relève par exemple que les délais imposés par les organisateurs sont difficiles, voire impossibles, à respecter dans un pays qui ne possède pas de service de courrier électronique ni de téléphone ou qui n'a pas les crédits nécessaires pour financer les déplacements dans les écoles. Parmi les autres problèmes, on peut citer l'absence de données précises sur les effectifs scolaires; la pénurie de compétences en matière de gestion; le peu de soin accordé aux détails, en particulier au niveau de l'édition, du codage et de la saisie des données; l'aide financière sommaire attribuée aux collaborateurs travaillant au projet; et la difficulté de garantir une bonne qualité d'impression des documents dans les délais impartis. Les consignes données aux surveillants des épreuves, comme celle d'aller et venir dans les allées des salles d'examen, sont manifestement hors de propos dans des salles de classe où il n'y a pas d'allée.

129. Malgré les nombreuses difficultés que soulèvent les évaluations internationales (voir *Encadré 9*), il peut être intéressant pour certains pays de disposer de données comparatives internationales. Il ne fait aucun doute qu'une étude comparative internationale peut fournir une multitude d'informations utiles sur la conception des épreuves, l'échantillonnage, l'analyse et la rédaction des rapports finaux. En outre, du fait du partage des coûts de conception et d'analyse, la charge incombant à chacun peut être réduite. Qui plus est, en Afrique du Sud, la pratique des évaluations internationales a eu des répercussions bénéfiques, telles que l'augmentation de la contribution du secteur privé à l'éducation, la fourniture de données de référence sur les niveaux de mathématiques et sciences et le renforcement des capacités en matière de recherche sur l'éducation. Par contre, rien ne permet d'affirmer que les résultats de l'évaluation ont entraîné un changement significatif de la politique éducative ou contribué au renforcement de capacités nationales durables en termes d'évaluation de l'éducation, puisque les membres des équipes TIMSS et TIMSS-R ont été transférés vers d'autres postes (Howie, 1999).

130. S'il est intéressant pour certains pays de disposer de données comparatives internationales, il convient d'attirer l'attention sur la nécessité d'en limiter la participation à des pays possédant des niveaux similaires de développement économique et social. A noter aussi que l'obtention de données comparatives internationales n'est pas un passage obligé pour bénéficier des nombreux avantages de la coopération internationale.

5. EVALUATIONS A L'ECHELON DE LA CLASSE

131. Evaluer l'apprentissage des élèves à l'échelon de la salle de classe (évaluation à la fois par les enseignants et par les élèves eux-mêmes) fait partie intégrante du processus d'enseignement-apprentissage. Cette forme d'évaluation revêt, pour une large part, un caractère subjectif, informel, immédiat, continu et intuitif, dans la mesure où elle interagit avec l'apprentissage au fur et à mesure de son avancement, le pilotage du comportement de l'élève, les performances scolaires et la réactivité à l'enseignement. Outre le suivi continu assuré par les enseignants, elle suppose de leur part de poser des questions et d'engager un dialogue pendant les cours, de corriger les devoirs faits à la maison et de constituer des dossiers. Sa fonction primaire est d'être un élément formateur. Pratiquée au cours de l'apprentissage (et non à la fin présumée de l'apprentissage), elle est destinée à améliorer l'acquisition de connaissances théoriques et pratiques par les élèves ou à les aider à les acquérir. Son but est de déterminer le niveau de connaissances théoriques, de connaissances pratiques ou de compréhension des élèves à un instant donné, de diagnostiquer les difficultés qu'ils sont susceptibles de rencontrer, de prendre des décisions quant aux prochaines étapes pédagogiques à accomplir (revoir certains points du programme ou poursuivre) et d'évaluer l'apprentissage qui s'est fait pendant un cours (voir *Encadré 10*).

Encadré 10. Evaluation à l'échelon de la classe

Une publication officielle britannique rappelle quelques principes de base à appliquer au titre d'une évaluation de l'enseignement et de l'apprentissage :

[Une évaluation] peut délimiter le cadre dans lequel seront définis les objectifs éducatifs et où seront reproduits et consignés les progrès des élèves. Elle peut servir de base à l'organisation des étapes pédagogiques ultérieures en fonction des besoins des élèves. En facilitant le dialogue entre les enseignants, elle peut développer les compétences professionnelles et aider l'école dans son ensemble à améliorer l'apprentissage au travers du programme scolaire ... (Grande Bretagne. Ministère de l'éducation et des sciences et ministère des Affaires galloises, 1988).

132. L'intérêt d'une évaluation réalisée par les enseignants est notamment qu'elle s'appuie sur une « évaluation des performances ». D'où plusieurs avantages : le contexte dans lequel se situent les connaissances théoriques et pratiques est préservé; elle donne des indications sur l'apprentissage de l'élève dans un environnement authentique; elle permet d'évaluer l'aptitude des élèves à exercer leur sens critique, à travailler en équipe, à résoudre des problèmes, à communiquer; enfin, elle peut faire considérablement progresser les élèves dans leur apprentissage et leur compréhension.

133. L'évaluation à l'échelon de la classe est parfois moins formelle qu'une interrogation ou un examen trimestriel. Plus objective, elle aura en outre une fonction cumulative, par exemple lorsque les informations qu'elle apporte sont utilisées pour décider du redoublement d'un élève ou de son passage dans la classe supérieure. Ces informations peuvent également être communiquées aux élèves, aux parents, ainsi qu'à d'autres enseignants ou personnes qui peuvent avoir besoin de connaître la progression d'un élève. De plus, elles peuvent servir à déterminer si les programmes scolaires, les méthodologies, l'organisation du travail en classe et les manuels scolaires sont bien appropriés et à ajuster, le cas échéant, les programmes scolaires en fonction des besoins et des aptitudes des élèves.

5.1. Qualité de l'évaluation à l'échelon de la classe

134. En dépit de son rôle central dans le processus d'enseignement-apprentissage, on ne dispose que de très peu d'informations sur la manière dont les enseignants évaluent leurs élèves. Il est néanmoins manifeste que la qualité des pratiques d'évaluation utilisées par les enseignants laisse parfois à désirer à de nombreux égards. Parmi les problèmes constatés, on peut citer la mauvaise formulation des questions, la prédominance de questions qui appellent des réponses courtes portant sur des connaissances factuelles, l'évocation de réponses qui font plus appel à la répétition qu'à la réflexion et l'absence de procédures conçues pour développer les facultés cognitives supérieures des élèves (Black & Wiliam, 1998; Madaus & Kellaghan, 1992).

135. Si l'on observe les pratiques employées dans les salles de classe en Afrique, on obtient un tableau relativement similaire. Et sans pour autant considérer qu'elles sont représentatives de toutes les écoles d'Afrique, il apparaît dans bien des cas, du moins implicitement, sinon explicitement, qu'elles sont valables pour un grand nombre d'écoles.

136. Considérons tout d'abord les observations concernant le système de questions-réponses qui est un élément important de l'évaluation. Une étude réalisée au Kenya montre que, dans de nombreux cours, les enseignants ne posaient aucune question aux élèves. Lorsqu'ils en posaient, il s'agissait de questions fermées - autrement dit, une formulation qui n'invite pas les élèves à développer des facultés cognitives plus élaborées. Qui plus est, la part laissée à l'évaluation de la compréhension des élèves avant de passer au chapitre suivant d'un cours est extrêmement réduite (Ackers et al, 2001).

137. Au Swaziland, les questions posées dans les classes de second cycle du secondaire sont, en grande majorité, de type rhétorique ou bien d'un faible niveau cognitif (Rollnick et al, 1998). Ce faible échelon taxinomique des questions a également été constaté dans des classes primaires en Tanzanie : les questions posées exigent des élèves qu'ils se rappellent des faits, individuellement ou en chœur (O-saki & Agu, 2002).

138. Les devoirs à la maison sont l'occasion pour les enseignants d'évaluer les compétences des élèves et d'obtenir un aperçu des difficultés éventuelles qu'ils peuvent rencontrer. Ceci dit, dans les écoles étudiées en Tanzanie, ces devoirs à la maison étaient relativement rares. Les raisons invoquées en l'occurrence étaient que, dans les écoles rurales, il n'y avait pas suffisamment de manuels scolaires ou de cahiers et que, dans les écoles urbaines, les gros effectifs des classes rendaient difficile la correction de nombreux devoirs. De plus, les enseignants corrigeaient rarement les cahiers des élèves, privant ces derniers des moyens de connaître leur niveau ou leurs progrès. Seuls quelques cahiers en effet contenaient des annotations des professeurs, annotations qui auraient pu être une manière de souligner un travail bien compris ou au contraire de repérer les lacunes d'un élève ayant rendu un mauvais devoir (O-saki & Agu, 2002).

139. D'aucuns attribuent les procédures d'évaluation qui prédominent dans les écoles au contexte d'enseignement-apprentissage lui-même, en l'occurrence un contexte où l'enseignant domine et où les élèves sont passifs. Selon certains observateurs, ce sont les enseignants qui parlent pendant toute la durée du cours, d'où le peu de place laissé pour la participation des élèves (Bude, 1993). D'après une étude réalisée à partir de l'observation des salles de classe et d'entretiens dans des écoles primaires en Tanzanie, les enseignants restent debout face à leur classe et exposent leur cours aux élèves (O-saki & Agu, 2002). Une étude sur l'interactivité en milieu scolaire réalisée sur 102 cours dans des écoles du Kenya aboutit au même constat : les enseignants adoptent une approche par « transmission », laquelle ne varie pratiquement pas, qu'il s'agisse d'un cours de mathématiques, d'anglais ou de sciences et qui ne laisse pratiquement aucune place à une participation active des élèves (Ackers et al, 2001). Les cours de sciences dans un lycée

au Swaziland sont également centrés sur l'enseignant : celui-ci pose des questions et les élèves répondent en chœur ou individuellement (Rollnick et al, 1998).

140. Parmi les autres raisons qui ont été avancées pour expliquer les pratiques d'évaluation existantes, on peut mentionner le manque de qualifications des enseignants, les effectifs élevés par classe, l'insuffisance de moyens, la pénurie de supports d'apprentissage (notamment des manuels scolaires) et l'absence d'endroits pour les stocker. Ainsi, en Guinée, les enseignants semblent peu formés aux techniques d'évaluation et la réalité est très « éloignée des procédures d'évaluation continue recommandées dans les programmes officiels » (Carron & Châu, 1996).

141. Autre point à souligner : l'effet négatif des examens nationaux sur les stratégies d'enseignement et d'évaluation mises en œuvre par les enseignants. Eisemon et al (1987) a ainsi constaté que l'enseignement consistait à faire des exercices, à réciter et à répondre à des questions du style « mots à compléter » et que toutes ces activités étaient destinées à inculquer aux élèves les techniques et connaissances factuelles dont ils auraient besoin pour passer l'examen public. Il semble par ailleurs que les examens « blancs » (conçus sur la base des examens nationaux) ont pour effet de réduire le champ de l'enseignement au domaine ou sujet d'examen correspondant à l'idée que s'en font les enseignants, d'où une certaine incohérence entre les cours traités. Bien qu'il soit difficile de définir les causes et les effets des abandons et des redoublements (voir N'tchougan6sonou, 2001), il est très probable que les pratiques des enseignants en matière d'évaluation de leurs élèves, ainsi que les examens nationaux, y contribuent.

5.2. Améliorer les évaluations à l'échelon de la classe

142. Si les observations exposées ci-dessus sur les évaluations à l'échelon de la classe sont révélatrices des pratiques courantes, ou les plus courantes, dans les écoles africaines, alors un profond changement s'impose, non pas seulement au plan des pratiques d'évaluation, mais aussi au plan de l'attitude globale des enseignants vis-à-vis de leurs méthodes pédagogiques, si l'on veut que l'évaluation puisse apporter une amélioration significative de l'apprentissage des élèves. Pour y parvenir, il faut que les enseignants intensifient leurs efforts pour solliciter la participation active des élèves pendant les cours, pour les impliquer davantage dans le processus d'enseignement-apprentissage et pour appeler leur attention sur la nécessité d'être co-acteur et co-auteur de leur propre apprentissage. D'un point de vue plus théorique, il importe que les enseignants prennent conscience du fait que l'apprentissage va au-delà de l'amélioration pure et simple des résultats aux examens et de l'acquisition pure et simple de connaissances et que les apprenants doivent non seulement acquérir des connaissances, mais aussi les générer, les maîtriser, les développer et les créer (Samoff, 1999). Toutes les initiatives susceptibles de faire progresser les systèmes éducatifs – formation préparatoire, formation continue, réforme des examens nationaux – dans cette voie sont à encourager.

143. Hormis quelques exceptions notables dans les pays francophones, l'attention portée à l'amélioration des évaluations à l'échelon de la salle de classe dans les réformes visant à utiliser les évaluations comme moyen d'améliorer l'apprentissage des élèves a été réduite. Comme on l'a évoqué précédemment, la difficulté de telles réformes et leur coût peuvent expliquer cette indifférence. Pourtant, étant donné son rôle central dans le processus d'enseignement-apprentissage, sans oublier les inconvénients que présentent les examens nationaux comportant un enjeu élevé, améliorer les pratiques d'évaluation à l'échelon de la salle de classe mériterait de constituer l'une des priorités de toute réforme visant à favoriser l'apprentissage des élèves (voir *Encadré 11*). De surcroît, il importe de

coordonner les initiatives destinées à améliorer les évaluations à l'échelon de la classe avec les réformes introduites dans d'autres systèmes d'évaluation.

144. Face à la complexité des évaluations à l'échelon de la classe et aux données recueillies sur les compétences et pratiques des enseignants en la matière, nécessité s'impose de développer une infrastructure propice à améliorer la qualité de ces évaluations. Cette infrastructure devrait a priori reposer sur la participation de tous les partenaires (tous ceux qui ont pouvoir d'influer sur les pratiques d'enseignement et d'évaluation dans les salles de classe, tels que les inspecteurs, concepteurs de programmes scolaires, les organisateurs d'examen et auteurs de manuels scolaires), sur la mise en place d'une formation préparatoire et d'une formation continue des enseignants et, autant que possible, sur la fourniture de documents d'évaluation aux écoles (épreuves et banques de questions par exemple). En Afrique du Sud par exemple, des banques de ressources d'évaluation (ARB, "Assessment Resource Banks") contenant des séries de devoirs conçus pour évaluer des résultats spécifiques d'apprentissage sont proposés aux écoles qui sont implantées dans des régions défavorisées et accueillent des élèves issus de communautés faibles du point de vue socioéconomique. Le principe de ces ARB consiste par exemple à archiver les progrès accomplis par les élèves dans des dossiers personnels et à transmettre périodiquement des rapports aux directeurs d'école qui, à leur tour, rendront compte périodiquement aux bureaux éducatifs du district (Kanjee, 2003). Au Swaziland, les documents fournis aux écoles contiennent des informations d'ordre général sur l'évaluation à l'échelon de la classe, des banques de questions, des épreuves, des spécifications d'examen et des supports pour résoudre des difficultés et améliorer la méthodologie (Mazibuko & Magagula, 2003).

145. Préparer les étudiants à évaluer leurs élèves devrait faire partie du programme de formation des futurs enseignants. On devrait leur présenter les concepts qui président à l'évaluation et leur enseigner comment utiliser, interpréter et juger de l'utilité de tout un ensemble de techniques formelles et informelles permettant d'évaluer les progrès effectués par leurs élèves et de diagnostiquer les difficultés qu'ils rencontrent. On devrait aussi leur enseigner les stratégies susceptibles de résoudre ces difficultés. Des programmes devraient être proposés dans lesquels les enseignants des écoles normales en psychologie, sciences de l'éducation, méthodologie des différentes matières et pédagogie travailleraient de concert.

5.3. Utilisation des évaluations à l'échelon de l'école dans des examens nationaux

146. Jusqu'à présent, notre étude a porté sur les évaluations à l'échelon de la salle de classe dans le contexte de l'enseignement quotidien, où les résultats de l'évaluation servent principalement à des fins formatrices. Venons-en maintenant à l'utilisation des résultats d'évaluation obtenus dans une salle de classe à des fins cumulatives (évaluations à l'échelon de l'école), où les résultats comptent dans la note attribuée aux élèves dans un examen public externe.

Encadré 11. Suggestions pour améliorer les évaluations à l'échelon de la salle de classe

On trouvera ci-dessous quelques suggestions destinées à améliorer les procédures d'évaluation appliquées par les enseignants à l'échelon de la salle de classe. Elles peuvent également servir de support pour la formation des maîtres.

- Les évaluations doivent être fréquentes, intégrées dans l'enseignement et comporter des questions axées sur des aspects essentiels de l'apprentissage.
- Les enseignants doivent fonder des espoirs raisonnables, mais néanmoins élevés, sur tous les élèves et recourir pour ce faire à une grande diversité de méthodes dans une grande diversité de situations (essais, devoirs à la maison et projets par ex.).
- L'évaluation doit mettre l'accent plus sur les éléments diagnostiques et formateurs que sur les éléments normatifs (c'est-à-dire que l'évaluation ne doit pas se contenter d'établir un classement des élèves en fonction de leurs résultats).
- Les enseignants doivent poser des questions qui permettent aux élèves de faire la preuve de facultés cognitives de haut niveau (et non pas seulement de la capacité de mémorisation) et qui exigent de raisonner par inférence et par déduction.
- Il y a lieu d'évaluer la compréhension qu'ont les élèves des principes généraux d'un domaine du programme scolaire, ainsi que leur aptitude à appliquer les méthodes et les stratégies de résolution de problèmes qui conviennent.
- Les résultats doivent être rendus rapidement aux élèves et accompagnés d'explications claires et précises.
- Il y a lieu d'évaluer les processus suivis par les élèves (la manière dont ils abordent et analysent une question) et non pas seulement les produits qu'ils ont rendus.
- L'évaluation doit aider les élèves à réfléchir à leur propre mode d'apprentissage.
- Les questions doivent exiger des élèves qu'ils approfondissent le sujet et en élargissent le champ, au lieu de se contenter de répéter des choses apprises.
- Les résultats des évaluations doivent être communiqués au moment opportun aux parents et aux autres personnes concernées (autres enseignants par ex.).
- L'utilisation d'épreuves établies sur la base de critères de référence peut améliorer les pratiques d'évaluation par les enseignants à l'échelon de la salle de classe. Des contrôles périodiques (à quelques semaines d'intervalle) permettront d'avoir un aperçu de l'apprentissage effectué par les élèves, de déterminer s'il y a lieu de revoir ou de poursuivre le programme et de repérer les élèves qui ont besoin de soutien complémentaire.

147. Parmi les critiques qui ont été faites des examens nationaux, on mentionnera qu'ils comportent des limites au niveau de la période sur laquelle ils se déroulent, des connaissances théoriques et pratiques qui sont évaluées et des techniques qui sont employées. De ce fait, les élèves n'ont pas toujours la possibilité de montrer leur véritable niveau de compétences.

148. Conscients de cette situation, plusieurs pays africains ont introduit, ou prévoient de le faire, un élément d'évaluation à l'échelon de l'école dans leurs examens nationaux. En effet, notamment dans l'enseignement primaire, d'aucuns espèrent à terme remplacer complètement les examens externes par des évaluations à l'échelon de l'école, considérant qu'elles sont la seule formule permettant d'évaluer valablement l'ensemble des compétences spécifiées dans les programmes scolaires et d'éviter les effets négatifs des examens externes sur l'enseignement et l'apprentissage. Ce système soulève toutefois des difficultés pour plusieurs raisons, notamment le risque qu'il ne soit pas perçu comme un outil « objectif » de sélection dans un pays où seuls quelques élèves ont accès aux écoles d'enseignement secondaire.

149. Plusieurs raisons ont été invoquées en faveur ou à l'encontre de l'utilisation des résultats d'évaluations réalisées à l'échelon de l'école pour déterminer les notes des élèves dans un examen qui est avant tout externe (Bude, 1997; Heyneman, 1988; Kellaghan & Greaney, 1992; Pennycuik, 1990a, 1990b). Commençons par les raisons avancées en faveur de la prise en compte des évaluations à l'échelon de l'école. Première raison : dans la mesure où une évaluation à l'échelon de l'école s'étend sur une période de temps relativement longue et qu'elle est réalisée par une personne qui connaît bien les élèves, l'appréciation qu'elle donne des résultats d'un élève sera vraisemblablement plus représentative et plus fiable que ne le permet un examen final externe unique. Deuxième raison : une évaluation à l'échelon de l'école permet d'élargir le champ des domaines du programme scolaire qui sont évalués. Un examen final ne permet pas d'évaluer de façon satisfaisante certains domaines d'aptitude des élèves, parmi lesquels leur capacité à préparer et organiser un projet, ainsi qu'à persévérer dans sa réalisation jusqu'à le mener à bonne fin. S'il est possible d'apprécier les capacités d'expression orale et les compétences pratiques des élèves dans un examen externe, cette appréciation demeure inévitablement limitée, artificielle et coûteuse.

150. Troisième raison : dans le cas d'une évaluation à l'échelon de l'école, les effets secondaires indésirables des examens externes sont limités, puisque les notes ne reposent pas exclusivement sur les performances des élèves à l'examen. Quatrième raison : une évaluation à l'échelon de l'école permet de compenser une mauvaise note inhabituelle d'un élève à un examen externe, pour cause de maladie par exemple. Cinquième raison : si l'évaluation à l'échelon de l'école est répartie sur l'année, elle peut renforcer le degré de motivation et d'efforts de l'élève pendant cette période.

151. A l'inverse, plusieurs raisons ont été invoquées pour démontrer les inconvénients potentiels d'une prise en compte des évaluations à l'échelon de l'école dans des examens de diplôme ou de sélection. Première raison : utiliser des évaluations faites à l'échelon de l'école pour délivrer des diplômes ou opérer une sélection risque de modifier la nature des relations entre les enseignants et les élèves, c'est-à-dire de privilégier le jugement critique au détriment du climat de soutien et de coopération. Deuxième raison : nombre d'enseignants sont considérés comme peu compétents pour mener à bien des évaluations et, quand bien même ils le sont, ils reconnaissent souvent qu'ils ne savent pas exactement comment exprimer ce jugement informel sous forme d'appréciations formelles et publiques et qu'ils ont alors recours à des formulaires rudimentaires ou à des questionnaires à choix multiple dénichés dans des manuels.

152. Troisième raison : les critères de notation appliqués dans le cadre d'une évaluation à l'échelon de l'école peuvent varier d'une école à l'autre, mais aussi d'une classe à l'autre dans une même école. Les enseignants ont tendance à évaluer leurs élèves par rapport à d'autres élèves de leur classe, et non par rapport aux élèves des autres écoles. Or, les examens nationaux sont censés fournir des données comparables entre écoles. Il faut par conséquent « pondérer » ou normaliser les résultats des évaluations à l'échelon de l'école par rapport aux résultats des examens écrits (comme en Namibie par

ex.). En réalité, cette méthode privilégie l'évaluation externe dans la mesure où les résultats sont ajustés par rapport aux notes obtenues à l'examen écrit et à leur distribution.

153. Quatrième raison : une évaluation organisée à l'échelon de l'école peut soumettre les enseignants à de fortes pressions de la part des parents, notamment dans des communautés très soudées. Cinquième raison : une évaluation à l'échelon de l'école prend un temps considérable aux enseignants qui doivent juger le niveau des élèves et tenir un registre des résultats. Pour certains enseignants, le surcroît de travail peut être trop important. Sixième raison : une évaluation à l'échelon de l'école pose quantité de problèmes administratifs aux écoles, concernant par exemple la façon de procéder en cas d'absence d'un élève aux épreuves ou en cas de transfert d'un élève d'une école à une autre. Septième raison : une évaluation réalisée par un enseignant est sujette à des dérives d'ordre sexiste ou liées aux antécédents socioéconomiques et aux caractères des individus. Enfin, il est difficile, voire impossible, d'intégrer dans ce type d'évaluation des candidats non rattachés à une école.

154. Au vu de ce qui précède, il n'est guère surprenant que la prise en compte d'évaluations à l'échelon de l'école dans des examens nationaux se soit révélée problématique dans plusieurs pays (Lesotho, Namibie, Nigeria, Ouganda, Swaziland, Tanzanie, par ex.). En dépit des efforts investis et d'une forte motivation en faveur de cette formule, les difficultés pratiques ont plus d'une fois conduit à abandonner le projet, à en différer l'exécution ou à réduire la part de cette évaluation à l'échelon de l'école à la portion congrue, sinon au rang de symbole. A titre d'exemple, la part de l'évaluation à l'échelon de l'école compte pour 5 % dans les notes des élèves à un examen public au Swaziland (l'objectif étant de la porter à 50 %) (Kellaghan, 2002). En Namibie par contre, la part de l'évaluation à l'échelon de l'école entre pour un tiers, voire la moitié des notes, selon la matière (V.d. Merwe, 1999).

155. Le statut privilégié accordé aux évaluations externes ne concerne pas uniquement la façon dont on pondère les évaluations à l'échelon de l'école par rapport à la part de l'examen externe. Face aux problèmes de compétences des enseignants en matière d'évaluation et à l'interprétation de leurs évaluations, certains pays qui ont intégré les évaluations à l'échelon de l'école dans leurs examens nationaux ont tendance à insister sur les aspects plus formels des évaluations à l'échelon de l'école au détriment des aspects informels. Les écoles reçoivent par exemple des épreuves écrites types ou des banques de questions auxquelles elles soumettent leurs élèves dans le cadre d'un examen formel, lequel n'est en fait qu'une simulation d'examen externe sous couvert d'évaluation à l'échelon de l'école. L'unique intérêt de cette pratique est qu'elle donne aux élèves une deuxième chance de montrer leurs aptitudes lors d'un examen écrit. Par contre, elle s'écarte de la finalité première d'une évaluation à l'échelon de l'école, qui est de donner l'occasion à des personnes (enseignants) qui connaissent les élèves d'apprécier leur véritable niveau de compétences dans différentes situations, par différents moyens, dans la durée et dans des domaines du programme scolaire qu'il est impossible de traiter dans un examen écrit.

6. FACTEURS SUSCEPTIBLES DE FAIRE OBSTACLE A L'UTILISATION DES RESULTATS D'EVALUATIONS ET D'EXAMENS A L'ECHELON DE LA SALLE DE CLASSE

156. Les efforts investis pour exploiter les résultats d'examens et d'évaluations nationales au niveau macro et faire remonter les conclusions ainsi obtenues sur le niveau scolaire des élèves jusqu'aux responsables de l'élaboration des politiques, planificateurs, gestionnaires, concepteurs de programmes scolaires, auteurs de manuels scolaires et simples citoyens sont fort louables. Mais il faut également assurer le retour de ces informations du niveau macro vers le niveau micro. S'il est couramment admis que « les réformes éducatives vivent ou meurent selon que leur mise en œuvre à l'échelon de l'école est une réussite ou un échec » (Verspoor, 1992, p. 23), rares sont les projets éducatifs qui prêtent suffisamment d'attention à la question de leur mise en œuvre. Cela peut être dû, en partie du moins, au fait qu'il est bien difficile de transmettre des informations aux enseignants et d'induire des changements dans leur comportement. L'espoir que les informations tirées des évaluations et des examens (via la publication des résultats, les bulletins d'informations, les ateliers) aboutissent à un changement radical de la culture scolaire et à une augmentation sensible du niveau de tous les élèves doit être tempéré par la prise en compte de divers facteurs susceptibles de rendre vaines les intentions des réformateurs. A défaut d'identifier ces facteurs et de les traiter, les politiques menées en matière d'évaluation risquent d'avoir un impact très limité sur la qualité d'apprentissage des élèves. Les facteurs analysés dans ce chapitre méritent d'être pris en considération dans tout projet de réforme, que ce soit des examens nationaux ou des évaluations nationales.

6.1. Mode d'intervention

157. Le mode d'intervention le plus couramment employé est fondé sur des structures administratives dans lesquelles les décisions concernant les intrants susceptibles de produire les comportements souhaités dans les écoles sont prises à l'échelon central. Cette approche est applicable lorsque les problèmes sont clairement structurés et appellent des solutions relativement standardisées. Elle l'est moins lorsque les problèmes ne sont pas clairement définis et que leur résolution passe par une multiplicité de réponses propres aux conditions locales. Il en va ainsi des problèmes qui portent sur des objectifs pédagogiques, sur l'équité et sur la qualité et qui, à ce titre, exigent des stratégies plus indirectes et interactives. Ces stratégies reposent en général sur des mesures administratives et législatives d'habilitation et d'encadrement, ainsi que sur un soutien financier et professionnel visant à prêter appui aux actions locales souhaitées (Verspoor, 1989, 1992).

6.1.1. Pertinence des informations

158. Il existe au moins deux limites inhérentes à l'utilisation des données dérivées des examens et évaluations nationales pour orienter les pratiques des enseignants. La première est que ces données, telles que celles qui sont contenues dans des rapports de jurys ou des rapports d'évaluation nationale, font référence à des niveaux généraux de performances du système éducatif. Par conséquent, elles ne reflètent pas forcément les conditions propres à une école déterminée ni le niveau des élèves dans cette école. La deuxième est que les questions posées dans les examens et les évaluations, de même que

les informations qui en découlent, ne peuvent donner qu'un aperçu limité des structures de connaissances ou des stratégies cognitives qu'elles sollicitent. Elles ne sont pas conçues pour permettre un diagnostic individuel de sorte que l'analyse des résultats ne peut donner sur chaque élève que peu d'informations réellement utilisables au plan pédagogique. La portée des conseils spécifiques qu'elles peuvent apporter aux enseignants n'en sera que plus limitée.

6.1.2. Compétences des enseignants

159. Imaginer que tous les enseignants dispensent un enseignement cohérent et structuré (c'est-à-dire conforme aux objectifs pédagogiques d'un examen) et sont capables d'évoluer et que les conditions de travail dans les salles de classe leur permettent de respecter les objectifs des réformes relève sans doute d'une conception irréaliste de l'ambiance d'une classe.

160. Comment espérer une adaptation des comportements dans la salle de classe, s'interrogent Carron et Châu (1996), alors que « la réalité tient davantage du chaos : les enseignants et les élèves sont souvent absents, l'enseignant n'a pas les qualifications requises pour exercer sa tâche correctement, il ou elle ne suit pas de plan de travail précis, se contente de lire ce qui est écrit dans les manuels, n'utilise pas le tableau noir... » (p. 202). Le manque de compétences des enseignants a également été constaté par d'autres observateurs. De nombreux enseignants ne maîtrisent pas bien la langue utilisée comme langue d'enseignement, ni la discipline qu'ils enseignent (sciences et agriculture notamment). En Afrique du Sud, près de la moitié des enseignants en mathématiques ne sont pas titulaires d'un diplôme officiel dans cette matière (Arnott & Kubeka, 1997). Selon une étude régionale MLA, moins d'un enseignant sur quatre au Malawi, à Madagascar et à l'île Maurice a suivi des études supérieures (Chinapah et al, 2000). Comme on l'a déjà indiqué, un rapport de jury d'examen pour l'épreuve de sciences appliquées du certificat de fin de premier cycle d'études secondaires au Swaziland montre clairement que, au vu des copies rendues par les candidats, les élèves de certains établissements ont reçu des informations factuelles erronées. Au Swaziland également, certains enseignants connaissent mal la langue employée comme langue d'enseignement. Ces observations concernent peut-être des cas extrêmes et exceptionnels, mais elles mettent en évidence des problèmes qui sont précisément ceux auxquels il convient de prêter attention dans le cadre d'une réforme.

6.1.3. Compréhension par l'enseignant des répercussions des changements

161. Le fossé qui existe entre ce qui est exigé dans une réforme d'une part et ce qu'en comprennent et font réellement les enseignants d'autre part peut être très profond. On ne peut pas s'attendre à ce que les enseignants connaissent le véritable enjeu d'un projet de réforme ou à ce qu'ils sachent quoi faire en fonction des résultats d'une évaluation. Même s'ils bénéficient d'une formation continue et de conseils, il peut y avoir des écarts non négligeables entre la signification attribuée à une réforme respectivement par les décideurs et par les enseignants; en effet, les enseignants interpréteront les conseils en s'appuyant sur leur compréhension du programme scolaire et sur leur pratique pédagogique, lesquelles peuvent être fort différentes de la conception qu'en ont les concepteurs des réformes (Chapman & Snyder, 2000, Grant, Peterson & Shojgreen-Downer, 1996).

162. En Ouganda par exemple, les enseignants ont eu d'énormes difficultés à comprendre les répercussions des modifications apportées aux examens sur leur pratique scolaire quotidienne et n'ont pas réussi à s'y adapter. Craignant de s'aventurer dans

l'application de stratégies pédagogiques nouvelles, ils ont préféré ne rien changer (Snyder et al, 1997).

163. Même constat pour Trinidad et Tobago. Une étude réalisée à partir de l'observation des salles de classe et d'entretiens avec des enseignants après une réforme des examens nationaux révèle que certains enseignants, qui mettaient l'accent sur la mémorisation et la récitation collective, n'ont pas su comment adapter leur enseignement pour développer les compétences des élèves en matière de résolution de problèmes. Ils se sont efforcés non sans mal de comprendre les répercussions de la réforme de l'examen sur leur pratique quotidienne; de leur côté, les élèves ont essayé de comprendre l'incidence éventuelle de la réforme à leur niveau (voir *Encadré 12*).

Encadré 12. Difficultés des enseignants à s'adapter à la réforme d'un examen

La principale conclusion qui ressort d'une étude réalisée à Trinidad et Tobago est que les enseignants n'avaient pas toujours les moyens de s'adapter aux changements apportés à un examen. Même s'ils comprenaient les exigences demandées dans l'examen du point de vue cognitif, ils étaient souvent incapables d'introduire les changements nécessaires à l'échelon de la classe afin d'améliorer les performances de leurs élèves en conséquence. De plus, les enseignants craignaient d'expérimenter des méthodes pédagogiques nouvelles du fait des pressions très fortes auxquelles ils étaient soumis, telles que la réputation d'être un bon professeur ou d'avoir traditionnellement de bons résultats aux examens locaux chez leurs élèves ... Pour bien des enseignants, le chemin menant à de meilleurs résultats aux examens semblait trop incertain et les risques d'adopter des techniques pédagogiques novatrices trop aléatoires.

(Chapman & Snyder, 2000, p. 362)

6.1.4. Complexité de l'enseignement

164. Il n'est pas besoin de souligner que l'enseignement est une activité extrêmement complexe où les enseignants sont en interaction constante avec un grand nombre d'apprenants, lesquels se situent à des degrés divers de niveau scolaire, d'aptitudes, de motivation et d'intérêt. Les règles qui régissent cette activité sont loin d'être simples, linéaires ou évidentes (voir *Encadré 13*).

Encadré 13. Complexité de l'enseignement

Les éléments clés d'un enseignement dans une salle de classe ne sont pas des comportements distincts et statiques, mais des modes de comportement dynamiques qui varient au gré des activités pratiquées et du climat général de la classe. Entre modifier des épreuves et modifier l'enseignement, le rapport n'est pas d'ordre technique. Changer la forme d'une épreuve, son contenu ou la complexité des questions posées n'est pas nécessairement signe d'un changement particulier de ce processus dynamique. De surcroît, les modifications de comportement d'un enseignant ne se traduisent pas nécessairement par des changements au niveau de l'apprentissage des élèves.

(Chapman & Snyder, 2000, pp. 462-463)

165. En outre, les enseignants doivent parfois recourir à un large éventail de stratégies pour résoudre les problèmes liés à de mauvais résultats scolaires. Même dans le cas d'élèves qui obtiennent des résultats identiques à un examen, la nature de ces résultats, les raisons qui en sont à l'origine et les véritables difficultés d'apprentissage qu'ils dénotent peuvent varier considérablement d'un élève à l'autre et, à ce titre, nécessiter de

recourir à des méthodes différentes pour les résoudre (voir par ex. Buly & Valencia, 2002).

166. Les problèmes peuvent venir du fait que, dans certains examens, comme c'est de plus en plus souvent le cas, l'accent est mis sur l'évaluation de compétences d'un niveau plus élaboré et non plus, ou plus exclusivement, sur des aptitudes générales de base. Il est à l'évidence très différent d'inculquer de tels comportements ou d'inculquer des comportements plus élémentaires d'apprentissage par cœur. Plus longues à enseigner, les compétences plus élaborées s'acquièrent progressivement avec le temps, sont moins sensibles aux méthodes pédagogiques directives, sont souvent difficiles à repérer dans un programme scolaire et sont parfois trop diffuses pour dicter le type d'enseignement (Airasian, 1988).

167. Cela ne veut pas dire qu'enseigner des compétences complexes soit une tâche impossible. Une étude menée au Nigeria sur des élèves de primaire auxquels les sciences humaines étaient enseignées selon une méthode de résolution de problèmes montre toutefois que non seulement ces élèves avaient acquis davantage de connaissances, mais aussi qu'ils en avaient mieux compris le contenu et avaient été capables d'appliquer ces connaissances à la résolution d'autres problèmes, ainsi que dans des évaluations (Ogundare, 1988).

6.1.5. Contexte de la salle de classe

168. S'agissant de réformes destinées à modifier le style d'enseignement, il importe de tenir compte du fait que de nombreux enseignants sont confrontés à des effectifs de classe très importants et ne disposent que de moyens limités (manuels scolaires et autres supports).

169. Du fait de la pénurie de manuels, il est rare que les élèves aient la possibilité d'étudier à partir de textes imprimés. Les enseignants présentent un résumé oral des textes, transformant des passages en exercices faits en classe et en devoirs à copier dans des cahiers (Eisemon et al, 1987). Au Kenya, des disciplines techniques (travaux manuels, arts ménagers) ont été ajoutées au programme scolaire de primaire en 1983, le but étant d'enseigner des techniques et des connaissances théoriques et pratiques considérées comme utiles pour le développement économique et de permettre aux élèves de suivre ultérieurement une formation professionnelle ou de se mettre à leur compte. Mais, en 1990, seuls 3.000 laboratoires sur les 13.000 requis avaient été construits (Rharade, 1997).

6.1.6. Résistance au changement à la perspective des risques qu'il peut présenter pour les élèves.

170. Aux yeux des parents et enseignants, changer les pratiques appliquées dans la salle de classe et, en particulier, les procédures d'évaluation risque de compromettre les avantages dont bénéficient les élèves dans le système existant (Chapman & Snyder, 2000). Bien que favorables à des mesures conçues pour promouvoir la qualité de l'éducation, ils peuvent être davantage soucieux de préserver l'avantage comparatif dont profitent leurs enfants avec le modèle actuel, par crainte que ces derniers ne soient moins performants dans un autre système ou n'éprouvent des difficultés pendant la phase d'adaptation à ce nouveau système.

7. CONCLUSION

171. L'amélioration de la qualité de l'éducation dans les pays en développement a été un élément clé des réformes qui ont été proposées ces dix dernières années par des gouvernements et des organismes bailleurs de fonds. Si le terme « qualité » est applicable à de multiples aspects de l'éducation, les résultats ou le niveau scolaire des élèves (acquisition de « connaissances théoriques et pratiques, de capacité de raisonnement et de valeurs utiles ») ont occupé une place centrale dans les analyses critiques de la qualité qui ont été faites au lendemain de la Déclaration de Jomtien sur l'Education pour tous en 1990.

172. Le présent document expose quatre catégories d'évaluation qui ont respectivement servi de toile de fond à l'élaboration de projets de réforme d'ampleur plus ou moins grande : examens nationaux, évaluations nationales, évaluations internationales et évaluations à l'échelon de la classe. Deux d'entre elles (examens nationaux et évaluations à l'échelon de la classe) sont depuis longtemps une caractéristique importante des systèmes éducatifs en Afrique. Toutes deux sont par ailleurs axées sur l'évaluation du niveau scolaire de chaque élève. En revanche, ce n'est qu'au cours de ces dix dernières années que les évaluations nationales et internationales ont pris leur place dans le paysage éducatif africain. D'autre part, contrairement aux examens nationaux et aux évaluations à l'échelon de la classe, leur finalité première est d'évaluer les performances du système éducatif, et non des élèves.

173. Les différences qui existent entre ces évaluations ont une incidence sur leur rôle potentiel dans l'amélioration de la qualité de l'éducation. En premier lieu, l'utilisation d'examens nationaux et d'évaluations à l'échelon de la classe comme vecteurs de réforme implique essentiellement de modifier les pratiques existantes, tandis que l'utilisation d'évaluations nationales (et internationales) exige la mise en œuvre de procédures novatrices. En deuxième lieu, la probabilité d'impact sur l'environnement d'apprentissage dans la salle de classe et sur le niveau scolaire des élèves est plus forte dans le cas d'examens nationaux et d'évaluations à l'échelon de la classe que dans le cas d'évaluations nationales (et internationales); en effet, ces dernières produisent plutôt des effets à distance dont la concrétisation passe par des décisions en matière de politique éducative et de gestion. En troisième lieu, les évaluations à l'échelon de la classe diffèrent des autres types d'évaluation de par leur caractère essentiellement « privé ». Les informations qu'elles fournissent ne sortent en général pas du cadre de la salle de classe, hormis lorsqu'elles sont communiquées aux parents ou, dans le cas de certains systèmes éducatifs, lorsqu'elles comptent dans la note attribuée aux élèves dans un examen public. Enfin, si l'évaluation à l'échelon de la classe est parfois la formule la plus efficace en termes d'impact sur les résultats scolaires des élèves, elle est aussi la plus difficile à modifier de l'extérieur.

174. Alors que l'objectif premier d'une évaluation de système ou d'une évaluation nationale est de refléter le niveau scolaire des élèves, pour les concepteurs de réforme qui souhaitent utiliser les évaluations pour améliorer la qualité de l'éducation, l'objectif est plus vaste. S'agissant des évaluations nationales, le but recherché est d'utiliser les informations ainsi obtenues, en s'appuyant sur une analyse des points faibles et des points forts relevés dans les connaissances théoriques et pratiques acquises par les élèves et sur les données relatives à la répartition des résultats au sein d'une population (par sexe, lieu, par ex.) comme outil d'aide à l'élaboration de politiques et à la prise de décisions (concernant l'allocation des ressources par exemple. Le but des évaluations nationales est également d'apporter des informations utiles aux concepteurs de programmes scolaires, aux auteurs de manuels scolaires, aux responsables de l'élaboration des politiques et plus généralement à l'opinion publique. Par la communication de ces informations, l'évaluation

pourrait faire office de levier des réformes tout en offrant un tableau représentatif de la situation d'un système éducatif.

175. Plus que d'autres types d'évaluation, néanmoins, ce sont les examens nationaux qui ont retenu l'attention dans les propositions visant à utiliser les évaluations comme levier de réformes pour améliorer le niveau scolaire des élèves. Il est admis depuis longtemps que le contenu et la forme des examens nationaux ont un impact immédiat et direct sur le contenu de l'enseignement et de l'apprentissage dans les écoles lorsque des enjeux importants sont liés aux performances des élèves (et des enseignants). Telle est la raison invoquée par plusieurs observateurs, au cours des vingt dernières années, pour défendre la thèse selon laquelle réformer les examens permettrait d'améliorer le niveau et la nature des performances scolaires des élèves.

176. Les réformes des examens qui sont proposées s'articulent autour de quatre lignes directrices. La première est d'améliorer la rentabilité administrative en introduisant, par exemple, l'usage de l'informatique et de la lecture optique. La seconde est d'accroître la qualité des examens en en supprimant les déficiences les plus patentes (comme le fait de ne couvrir qu'une partie du programme scolaire ou de privilégier la mémorisation et la reconnaissance). Une troisième réforme vise à inclure dans les examens des thèmes censés être utiles aux nombreux élèves qui ne poursuivront pas d'études secondaires et, par là même, à éviter le risque de dérive des examens vers leur fonction de sélection. Une quatrième réforme concerne la modification du contenu des questions et de leur échelon taxinomique en y ajoutant des éléments qui permettent de mesurer les aptitudes intellectuelles plus élaborées et la capacité à appliquer des connaissances théoriques et pratiques dans des situations nouvelles; cette réforme serait bénéfique aussi bien pour les élèves qui quittent l'école après avoir passé l'examen que pour ceux qui poursuivent une éducation formelle. Compte tenu du fait que les enseignants « cadrent leur enseignement avec l'examen », l'application de ces réformes devrait en principe se refléter dans le contenu et la méthode d'enseignement, lesquels devraient à leur tour se refléter dans les résultats scolaires des élèves. Pour en renforcer l'impact, il peut s'avérer judicieux de prévoir un retour d'informations vers les écoles, c'est-à-dire des conclusions tirées de l'analyse des performances des candidats aux examens dans les domaines où des difficultés ont été constatées.

177. L'intérêt de ces propositions ne fait aucun doute et les pays auraient tous grand avantage à les appliquer dans leurs systèmes d'examen. Par ailleurs, dans bien des cas, le surcoût de ces indispensables réformes est faible, voire nul. De multiples possibilités existent pour les mettre en œuvre, notamment le jumelage avec des organismes outre-mer aux fins d'obtenir une assistance technique ou le renforcement de relations professionnelles entre des jurys d'examens, tel qu'il en existe déjà dans certaines régions d'Afrique, afin de faciliter les échanges de personnel et d'améliorer le pilotage des normes. D'autres initiatives pourraient être envisagées au titre d'un renforcement des capacités régionales, comme la recherche de solutions pour garantir le maintien des niveaux scolaires, lutter contre les pratiques frauduleuses et offrir une formation sur de nouvelles techniques d'évaluation et de notation.

178. On peut s'attendre à ce que la diversité des réformes et des innovations introduites aura pour effet d'améliorer la qualité des examens. Pour autant, il ne faudrait pas imaginer que l'amélioration d'une évaluation se traduira automatiquement par une couverture plus large des programmes scolaires et un meilleur niveau scolaire des élèves du seul fait que la proposition inverse est vraie, c'est-à-dire qu'une évaluation de mauvaise qualité tend à restreindre le champ du programme scolaire traité et à abaisser le niveau des élèves (Torrance, 1997). Au vu des données disponibles, tout porte à croire que, si l'on modifie le contenu d'une matière présentée à un examen (introduction d'un nouveau sujet ou d'un nouvel élément d'un sujet, comme la rédaction d'un essai, dans un examen

par ex.), le contenu qui sera enseigné en classe aux élèves sera modifié en conséquence. Par contre, pour ce qui est du niveau général de performances des élèves ou de leurs facultés cognitives, les données sont plus floues, mais on peut néanmoins escompter une légère amélioration. Une chose est sûre : un grand nombre d'élèves aujourd'hui ont des résultats médiocres aux examens. Ce n'est pas tant que leur champ de performances est limité ou que le niveau taxinomique de leurs processus cognitifs est faible. La cause ne peut en être imputée à la qualité des examens, car leur incidence en l'occurrence est vraisemblablement moindre que les compétences des enseignants, les effectifs élevés, la pénurie de moyens et de ressources et la difficulté d'inculquer des compétences plus élaborées. Il n'est pas réaliste d'imaginer que la réforme d'un examen aura plus d'effets que ces autres facteurs.

179. Le fait que certaines réformes d'examen soient dictées par des mécanismes où l'enjeu est important (sélection, compétition, publication des résultats par ex.) incite à s'interroger sur les effets négatifs (s'ils ne sont pas intentionnels) qui peuvent s'ensuivre. S'il est vrai que les examens comportant un enjeu élevé ont toutes les chances d'influer sur l'enseignement et l'apprentissage, le risque est que les principaux bénéficiaires en soient les bons élèves. Tel sera le cas par exemple si les enseignants, davantage préoccupés par le taux de réussite de leurs élèves, portent toute leur attention sur les élèves qui offrent le meilleur potentiel de réussite; si l'on fait redoubler les élèves médiocres pour les empêcher de passer l'examen; et si des élèves, découragés, quittent l'école avant d'avoir achevé leur scolarité primaire. Ce sont là des questions importantes à considérer dans le contexte de la politique EPT qui a, comme objectif, « d'obtenir pour tous des résultats d'apprentissage reconnus et quantifiables, notamment en ce qui concerne la lecture, l'écriture, le calcul, ainsi que les compétences indispensables dans la vie courante ». Il est essentiel de prendre toute la dimension du conflit qui existe entre la fonction traditionnelle des examens d'une part, qui était la sélection des élèves les plus performants et, partant, les plus aptes à poursuivre leurs études, et les objectifs de l'EPT d'autre part, qui soulignent l'attention qui doit être portée à *tous* les élèves. De même, il est nécessaire d'adopter les mesures qui s'imposent pour que les examens ne soient pas un obstacle à la réalisation des objectifs de l'EPT. L'une des premières mesures à prendre pourrait consister à intégrer dans les examens, notamment les examens de fin de cycle primaire, un contenu qui corresponde à tous les niveaux de performances des élèves.

180. Durant les années 90 en Afrique, de nombreuses initiatives ont été lancées dans le domaine des évaluations nationales. Il semble que l'impulsion et les ressources octroyées par des organisations internationales et d'autres bailleurs de fonds au lendemain de Jomtien en aient été le moteur. Dans cette hypothèse, on peut alors se demander si les décideurs et les planificateurs ont véritablement ressenti dans leur propre pays la nécessité d'effectuer des évaluations et s'ils se sont fixés des échéances pour leur réalisation. Le fait que plusieurs pays aient réalisé trois évaluations nationales indépendantes (respectivement parrainées ou financées par un organisme différent) laisse supposer que les évaluations ne sont pas une réponse à des besoins perçus à l'échelon local et qu'elles ne font pas partie intégrante des structures et activités normales des ministères. Compte tenu du coût de ces opérations et des problèmes relatifs au recrutement de personnel compétent pour les mener à bien, l'urgence d'une rationalisation se fait clairement sentir.

181. Toutes les évaluations nationales menées en Afrique relèvent du même schéma : les évaluations des performances des élèves portent sur un échantillon d'écoles et sur des matières fondamentales du programme et elles intègrent des informations sur les antécédents des enseignants et des élèves, lesquelles sont par la suite mises en corrélation avec les performances des élèves. (Quelques évaluations diffèrent de ce schéma type; pour certaines études, des tests étaient prévus pour les enseignants.) Le modèle entrées-sorties utilisé dans cette approche qui s'inscrit dans le droit fil des

premières évaluations internationales des performances (voir Husén & Postlethwaite, 1996) a eu tendance à privilégier la fonction recherche des évaluations (c'est-à-dire identifier ou « comprendre » les causes qui expliquent les performances scolaires) au détriment de leur fonction gestion (c'est-à-dire fournir aux décideurs des informations en retour sur l'état du système éducatif). Il a également incité à rechercher l'existence éventuelle de liens de causalité entre les facteurs intrants et les résultats scolaires des élèves. Ce type de démarche peut aboutir à des conclusions hasardeuses, voire erronées, pour trois raisons au moins. La première raison est que, d'une manière générale, il n'est pas possible d'établir de rapports de causalité à partir de données issues d'une étude transversale. La seconde est que, compte tenu des instruments employés pour collecter des informations sur les antécédents et de la manière dont les variables d'analyse sont construites, il n'est pas toujours possible d'identifier les principaux aspects de l'environnement qui sont liés aux résultats (par ex. le capital « culturel » et « social » du foyer et de la communauté plutôt que le statut socioéconomique de la famille). Troisième raison enfin, les personnes chargées de l'organisation des évaluations nationales n'ont pas nécessairement l'expérience de l'instrument d'analyse statistique (modélisation linéaire hiérarchique par ex.) auquel elles doivent avoir recours pour explorer ces corrélations. C'est pourquoi il peut s'avérer plus opportun de considérer les données fournies par une évaluation nationale comme un outil de gestion stratégique utilisable dans le cadre de la planification de l'éducation (Verspoor, 1992). Une évaluation serait perçue et traitée dans une perspective « gestionnaire » en tant qu'aide décisionnelle (concernant par ex. la nature de la formation continue requise) et les études comportant une recherche plus complexe des « causes » des performances scolaires relèveraient de formes de recherche traditionnelles.

182. Ainsi qu'on a pu le constater, les conclusions des évaluations nationales sont prises en compte dans le cadre de délibérations des pouvoirs publics, d'études de synthèse et d'ouvrages, même si la réalité varie d'un pays à l'autre. On peut supposer que c'est plus particulièrement le cas dans les pays où les hauts responsables et les gestionnaires ont contribué à l'instauration d'une évaluation nationale et au choix du public visé. L'impact des conclusions des évaluations nationales sur les pratiques scolaires est moins facile à percevoir. Il n'y a rien d'étonnant à cela. Les évaluations nationales sont relativement récentes et on peut comprendre aisément qu'il faille du temps avant d'en ressentir les effets. Qui plus est, obtenir des résultats qui influent sur les pratiques scolaires n'a rien d'une tâche facile. Malgré les efforts investis dans plusieurs pays, leur impact n'a pas encore été évalué. On ignore encore comment s'y prendre au mieux en la matière.

183. En ce qui concerne le rôle des évaluations nationales, il ne faut pas oublier que leur succès requiert un haut niveau de compétences techniques. Il est essentiel de renforcer les capacités pour ce faire si l'on veut doter les décideurs et les gestionnaires d'informations valables, fiables et utiles à la prise de décision. Si la qualité de conception, d'instrumentation, d'échantillonnage et d'analyse fait défaut, le jeu n'en vaudra peut-être pas la chandelle. Pour qu'il en soit ainsi, il est indispensable que les pouvoirs publics et les bailleurs de fonds prévoient l'octroi d'un budget suffisant.

184. Un petit nombre seulement de pays africains a participé aux évaluations internationales destinées aux pays industrialisés, ce qui est tout à fait compréhensible et justifié. Toutefois, quelques évaluations nationales organisées en Afrique autorisent une comparaison de performances d'un pays à l'autre. En outre, certains pays semblent intéressés par l'idée de développer leurs capacités dans ce domaine. Pour autant, le principal intérêt d'une coopération internationale réside dans la mise en commun de ressources en vue de développer les compétences nécessaires pour mener à bien des évaluations à l'échelon national et non pas de comparer les performances des pays entre eux.

185. Comme on l'a indiqué précédemment, les pratiques d'évaluation à l'échelon de la classe appliquées par les enseignants ont toutes les chances d'améliorer le niveau scolaire des élèves. De même, on a évoqué le fait que ces pratiques d'évaluation étaient médiocres et, en particulier, peu propices à promouvoir le développement de compétences intellectuelles plus élaborées et d'aptitudes à la résolution de problèmes chez les élèves. Il est malheureusement plus difficile et plus coûteux d'améliorer les pratiques d'évaluation des enseignants que d'améliorer ou de concevoir d'autres formes d'évaluation. Améliorer la formation initiale des enseignants et mettre en place une formation continue spécifique nécessitent inévitablement des ressources supplémentaires. Il est indispensable que les nombreux enseignants qui ont été embauchés à la hâte pour satisfaire aux exigences du programme EPT bénéficient d'une formation sur les évaluations à l'échelon de la classe.

186. Pour compenser quelques-uns des inconvénients des examens nationaux, qui sont imputables au champ limité de connaissances théoriques et pratiques ainsi évaluées et aux enjeux importants liés aux performances des élèves, une solution peut consister à intégrer les évaluations des enseignants dans les notes d'examen attribuées aux élèves. Si cette pratique existe déjà ou est en passe de l'être dans plusieurs systèmes d'examen, elle n'en soulève pas moins des difficultés. Il est donc probable que les évaluations à l'échelon de l'école ne parviendront pas à remplacer complètement les examens externes de sélection tant que le nombre de places disponibles dans l'enseignement au-delà du primaire sera aussi limité qu'il l'est aujourd'hui dans maints pays africains. La prise en compte des évaluations à l'échelon de l'école dans les notes d'examen comporte pourtant des avantages incontestables. La formation continue et les supports dont ont besoin les enseignants pour justifier l'intégration des évaluations à l'échelon de l'école dans des examens devraient avoir un effet bénéfique sur leurs pratiques d'évaluation dans leur ensemble. Ces dernières devraient à leur tour avoir un impact sur les élèves les plus faibles.

187. L'une des conclusions de l'analyse présentée dans ce document est que, si les résultats des évaluations *peuvent* améliorer la politique et la gestion des ressources éducatives et peuvent façonner les pratiques pédagogiques des enseignants, le succès n'est pas assuré. Ce succès dépendra tout d'abord de la volonté politique des gouvernements de soutenir ces efforts. Ces efforts devront s'exercer à deux niveaux. En premier lieu, il conviendra d'institutionnaliser les systèmes d'évaluation nationale et de les intégrer dans les structures et processus d'élaboration des politiques et de prise de décision des dirigeants. Il faudra, pour ce faire, intégrer les informations tirées des évaluations dans des Systèmes de gestion de l'information en éducation (EMIS) et instaurer des procédures aptes à garantir que ces informations seront accessibles à tous les acteurs participant à l'élaboration des politiques et à la gestion. Pour ce qui est du renforcement des capacités requises en la matière, les pays pourront recourir aux Systèmes d'informations statistiques nationales sur l'éducation (NESIS, "National Education Statistical Information Systems"), un programme lancé par le Groupe de travail sur les statistiques éducatives (WGES, "Working Group on Education Statistics") de l'Association pour le développement de l'éducation en Afrique (ADEA). (A ce jour, 47 pays ont participé au programme NESIS). Il importera également de veiller à aligner les évaluations nationales sur d'autres grands principes d'orientation pédagogique du système éducatif – autres systèmes d'évaluation (y compris l'alignement des normes de niveau), programmes scolaires, formation des maîtres, renforcement des capacités des écoles et mesures visant à supprimer les injustices.

188. Des actions devront en outre être engagées à l'échelon de l'école. Il conviendra d'ouvrir des circuits de communication adaptés pour informer les enseignants des changements opérés. Dans certains cas, ce sera sans doute relativement facile. S'agissant de problèmes d'espace ou de fournitures par exemple, des actions directes et relativement simples à mettre en œuvre seront instaurées. Par contre, dans le cas de

problèmes touchant aux objectifs pédagogiques, à l'équité et à la qualité, la tâche risque d'être difficile, et les enseignants pourront avoir besoin d'un soutien considérable et durable pour interpréter les réformes et concevoir des stratégies d'enseignement adaptées. Inspecteurs, surveillants et directeurs d'établissement, qui tous auront peut-être besoin d'une formation, ont un rôle à jouer en la matière.

189. Nombreux sont ceux qui, à la lumière de leurs observations sur l'éducation en Afrique, font état du succès très mitigé des réformes éducatives engagées par le passé. Il est vrai, semble-t-il, que les efforts isolés et ponctuels n'ont, dans l'ensemble, pas été à la hauteur des attentes. Les propositions qui visent à utiliser les évaluations pour améliorer la qualité de l'éducation auront-elles plus de succès? La réponse sera négative si les réformes des évaluations ne sont pas intégrées dans les structures et processus d'élaboration des politiques et de prise de décision et ne sont pas alignées sur d'autres grands systèmes d'orientation pédagogique. A l'inverse, si elles le sont, l'optimisme sera de mise.

190. Il existe d'autres raisons d'être optimiste en dépit des multiples difficultés que rencontre le secteur éducatif en Afrique, ainsi qu'on l'a évoqué au début de ce document. Au lendemain de Dakar, tant dans la formulation des Objectifs de développement du millénaire (MDG, "Millennium Development Goals") que dans les priorités énoncées dans les Documents de stratégie pour la réduction de la pauvreté (PRSP, "Poverty Reduction Strategy Papers"), les politiques, les plans d'application et les diverses initiatives, se dégage une volonté de conférer un sentiment d'urgence et de créer un climat de transparence. Réformer les modes d'évaluation devrait permettre d'obtenir des données objectives qui n'étaient pas disponibles auparavant et qui seraient utiles pour la prise de décision. Les efforts investis dans la réforme des évaluations bénéficieront par ailleurs du soutien des partenaires de l'ADEA qui se mobilisent en faveur du développement d'une « culture » de recherche de solutions et de moyens d'action dans le contexte africain pour résoudre les problèmes et éliminer les obstacles (ADEA, 1999). Si ces conditions sont réunies, on peut penser que les évaluations joueront un rôle important dans l'amélioration de la qualité de l'apprentissage des élèves.

7.1. Recommandations

191. Quelques recommandations sont proposées, ci-après; elles pourraient être d'un certain intérêt pour les ministères de l'éducation, les divers responsables en matière d'éducation et autres parties prenantes:

- Les ministères de l'éducation, en collaboration avec tous ses partenaires, devraient élaborer des politiques globales en matière d'évaluation, dans lesquelles les objectifs d'un système d'évaluation (examens nationaux, évaluations nationales et évaluation à l'échelon de la classe) seraient indiqués.
- Les systèmes d'évaluation seraient conçus pour être complémentaires dans l'amélioration de l'apprentissage des élèves.
- Les politiques d'évaluation devraient tenir compte des politiques qui administrent d'autres volets du système éducatif et devraient les appuyer (cela comprend la formation des maîtres, le renforcement des capacités des écoles, les mesures pour remédier aux injustices) afin de fournir des approches cohérentes, systémiques et variées afin d'améliorer l'apprentissage des élèves.
- Les informations tirées des évaluations devraient être intégrées dans les structures d'élaboration des politiques et de prise de décision des ministères de l'éducation, afin de servir d'intrant dans la planification de l'éducation et la gestion stratégique; elles devraient aussi être transmises aux responsables des programmes scolaires et des manuels.

- Les curricula et les examens nationaux devraient être révisés afin de les adapter à la vie quotidienne des élèves et aux besoins de l'économie et il faudrait procéder aux ajustements nécessaires. Cet exercice devrait avoir toujours à l'esprit les objectifs de l'Education pour tous.
- Il conviendrait de poursuivre les efforts pour réformer les examens nationaux afin d'améliorer la couverture du curriculum et la qualité de l'apprentissage des élèves.
- Les examens nationaux devraient être révisés afin de vraiment refléter leur fonction de certification tel que cela apparaît dans leur contenu et leur niveau de difficulté; des ajustements seront faits si cela est nécessaire.
- L'éventail des examens nationaux devrait être élargi afin que les diplômes contiennent des informations sur des aspects du développement de l'élève que l'examen écrit de fin de cycle ne peut pas mesurer.
- Les évaluations nationales devraient répondre aux besoins d'informations des ministères de l'éducation et des autres parties prenantes. Il faudrait rationaliser les activités d'évaluation dans les pays ayant plusieurs systèmes éducatifs.
- Les ministères de l'éducation devraient s'efforcer d'atteindre des normes techniques élevées dans les évaluations nationales en assurant le développement des moyens techniques et des capacités organisationnelles afin d'assurer le suivi des performances des élèves.
- Pour reconnaître le rôle central joué par les enseignants dans l'amélioration de la qualité d'apprentissage des élèves, il conviendrait de :
 - Améliorer les compétences des enseignants en matière d'évaluation à la fois dans la formation initiale et la formation continue, qui doivent tenir compte des conditions dans lesquelles les enseignants travaillent (taille de la classe, fourniture des ressources nécessaires);
 - Donner des conseils pour utiliser les informations données par l'évaluation afin de procéder aux décisions de passage dans la classe supérieure;
 - Fournir des exemples de bonnes pratiques d'évaluation à la fin des manuels scolaires;
 - Développer les moyens pour transmettre aux enseignants les informations sur les évaluations et les réformes, et donner des conseils pour interpréter le sens des informations pour la pratique pédagogique.
- Il conviendrait de travailler à l'établissement d'équivalences entre les résultats/qualifications attribués dans le système formel et le système informel du secteur de l'éducation.

8. REFERENCES BIBLIOGRAPHIQUES

- Ackers, J., Migoli, J., & Nzomo, J. (2001). Identifying and addressing the causes of declining participation rates in Kenyan primary schools. *International Journal of Educational Development*, 21, 361-374.
- ADEA (Association for the Development of Education in Africa). (1999). *What works and what's new in education: Africa speaks*. Paris: Author.
- ADEA. (2002). Reaching out, reaching all. Sustaining effective policy and practice for education in Africa. Fighting HIV/AIDS. Proceedings of the ADEA Biennial Meeting, Arusha, Tanzania, October 7-11, 2001. Paris: Association for the Development of Education in Africa.
- Airasian, P.W. (1988). Measurement driven instruction: A closer look. *Educational Measurement: Issues and Practice*, 7, 6-11.
- Arnott, A., & Kubeka, Z. (1997). *Mathematics and science teachers. Demand, utilization, supply and training in South Africa*. Johannesburg: Edusource.
- Bamgbose, A. (1991). *Language and the nation: The language question in Sub-Saharan Africa*. Edinburgh: University Press.
- Beaton, A.E., Postlethwaite, T.N., Ross, K.N., Spearritt, D., & Wolf, R.M. (1999). *The benefits and limitations of international achievement studies*. Paris: UNESCO: International Institute for Educational Planning; International Academy of Education.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Bude, U. (1993). Strategies for using information to improve learning conditions and instructional practices at the school level. In D.W. Chapman & L.O. Mählck (Eds.), *From data to action: Information systems in educational planning* (pp. 68-91). Paris: UNESCO: International Institute for Educational Planning.
- Bude, U. (1997). End of primary school examination in eastern and southern Africa: An overview. In U. Bude & K. Lewin (Eds.), *Improving test design. Vol 2. Assessment of science and agriculture in primary schools in Africa; 12 country cases reviewed* (pp. 12-30). Bonn: Education, Science and Documentation Centre.
- Bude, U., & Lewin, K. (1997). Introduction. In U. Bude & K. Lewin (Eds.), *Improving test design. Vol 1. Constructing test instruments, analysing results and improving assessment quality in primary schools in Africa* (pp. 6-11). Bonn: Education, Science and Documentation Centre.
- Buly, M.R., & Valencia, S.W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis*, 24, 219-239.
- Burnstein, L. (Ed.). (1993). *The IEA study of mathematics III: Student growth and classroom processes*. Oxford: Pergamon.
- Carron, G., & Châu, T.N. (1996). *The quality of primary schools in different development contexts*. Paris: UNESCO Publishing/International Institute for Educational Planning.
- Chapman, D.W., & Mählck, L.O. (1993). Improving educational quality through better use of information. In D.W. Chapman & L.O. Mählck (Eds.), *From data to action: Information systems in educational planning* (pp. 1-22). Paris: UNESCO: International Institute for Educational Planning.
- Chapman, D.W., & Snyder, C.W. (2000). Can high stakes national testing improve instruction?: Reexamining conventional wisdom. *International Journal of Educational Development*, 20, 457-474.
- Chinapah, V. (1997). *Handbook on monitoring learning achievement. Towards capacity building*. Paris: UNESCO.
- Chinapah, V. et al. (2000). *With Africa for Africa. Towards quality education for all*. Pretoria: Human Sciences Research Council.

- Cizek, G.J. (2001). *Setting performance standards. Concepts, methods, and perspectives*. Mahwah NJ: Lawrence Erlbaum.
- Clayton, T. (1998). Explanations for the use of languages of wider communication in education in developing countries. *International Journal of Educational Development*, 18, 145-157.
- Eisemon, T.O. (1990). Examination policies to strengthen primary schooling in African countries. *International Journal of Educational Development*, 10, 69-82.
- Eisemon, T.O., Patel, V.L, & Abagi, J. (1987). Read these instructions carefully: Examination reform and improving health education in Kenya. *International Journal of Educational Development*, 8, 55-66.
- ERGESE (Evaluative Research of the General Education System in Ethiopia). (1986). *A quality study: Summary report presented to the Executive Committee of ERGESE*. Addis Ababa: Author.
- Eritrea. Department of General Education. (1999). *Competence test report 1998/99*. Asmara: Author.
- Grant, S., Peterson, P., & Shojgreen-Downer, A. (1996). Learning to teach mathematics in the context of system reform. *American Educational Research Journal*, 33, 500-543.
- Greaney, V., & Kellaghan, T. (1996a). *Monitoring the learning outcomes of education systems*. Washington DC: World Bank.
- Greaney, V., & Kellaghan, T. (1996b). The integrity of public examinations in developing countries. In H. Goldstein & T. Lewis (Eds.), *Assessment: Problems, developments and statistical issues* (pp. 167-188). New York: Wiley.
- Great Britain. Department of Education and Welsh Office. (1988). *Task Group on Assessment and Testing. A report*. London: Her Majesty's Stationery Office.
- Grisay, A., & Mählick, L. (1991). *The quality of education in developing countries: A review of some research studies and policy documents*. Paris: International Institute for Educational Planning.
- Heyneman, S.P. (1988). Improving university selection, educational research, and educational management in developing countries: The role of examinations and standardized testing. In S. Heyneman & I. Fägerlind (Eds.), *University examinations and standardized testing* (pp. 197-216). Washington DC: World Bank.
- Heyneman, S.P., & Ransom, A.W. (1992). Using examinations and testing to improve educational quality. In M.A. Eckstein & H.J. Noah (Eds.), *Examinations: Comparative and international studies* (pp. 105-120). Oxford: Pergamon.
- Himmel, E. (1996). National assessment in Chile. In P. Murphy et al (Eds.), *National assessments: Testing the system* (pp.111-128). Washington DC: World Bank.
- Howie, S.J. (1999). *National assessment in an international context: Building capacity and expertise in South Africa*. Paper presented at World Bank Conference on Human Development, March 3-5.
- Howie, S.J. (2001). *Mathematics and science programmes in grade 8 in South Africa 1989/99. TIMSS-R South Africa*. Pretoria: Human Sciences Research Council.
- Howie, S.J. (2002). *English language proficiency and contextual factors influencing mathematics achievement of secondary school pupils in South Africa*. PhD dissertation, University of Twente.
- Husén, T., & Postlethwaite, T.N. (1996). A brief history of the International Association for the Evaluation of Educational Achievement (IEA). *Assessment in Education*, 3, 129-141.
- Kanjee, A. (2003). Using assessment resource banks to improve the teaching and learning process. In *Improving the quality of primary education: Good practices and emerging models of district development* (pp. 59-71). Pretoria: District Development Support Programme/Research Triangle Institute.

- Kellaghan, T. (1992). Examination systems in Africa: Between internationalization and indigenization. In M.A. Eckstein & H.J. Noah (Eds.), *Examinations: Comparative and international studies* (pp. 95-104). Oxford: Pergamon.
- Kellaghan, T. (1996a). Can public examinations be used to provide information for national assessment? In P. Murphy, V. Greaney, M.E. Lockheed, & C. Rojas (Eds.), *National assessments: Testing the system* (pp. 33-48). Washington DC: World Bank.
- Kellaghan, T. (1996b). IEA studies and educational policy. *Assessment in Education*, 3, 143-160.
- Kellaghan, T. (2002). *Assessment in the Swazi system of education*. Dublin: Educational Research Centre.
- Kellaghan, T. (2003). Local, national, and international levels of system evaluation: Introduction. In T. Kellaghan & D.L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 873-882). Dordrecht: Kluwer Academic.
- Kellaghan, T., & Farrell, E. (1998). *A study of the feasibility of obtaining, and the utility for describing learning outcomes of, data relating to public examination performance at the end of primary schooling*. Report submitted to UNESCO on behalf of the International Association for Educational Achievement.
- Kellaghan, T., & Greaney, V. (1992). *Using examinations to improve education. A study in fourteen African countries*. Washington DC: World Bank.
- Kellaghan, T., & Greaney, V. (2001a). The globalisation of assessment in the 20th century. *Assessment in Education*, 8, 87-102.
- Kellaghan, T., & Greaney, V. (2001b). *Using assessment to improve the quality of education*. Paris: UNESCO: International Institute for Educational Planning.
- Kellaghan, T., & Madaus, G.F. (2000). Outcome evaluation. In D.L. Stufflebeam, G.F. Madaus & T. Kellaghan (Eds.), *Evaluation models. Viewpoints on educational and human services evaluation* (2nd ed., pp. 97-112). Boston: Kluwer Academic.
- Kellaghan, T., Madaus, G.F., & Raczek, A. (1996). *The use of external examinations to improve student motivation*. Washington DC: American Educational Research Association.
- Kelly, M.J. (1991). *Education in a declining economy. The case of Zambia, 1975-1985*. Washington DC: World Bank.
- Kudjoh, A., & Mingat, A. (1993). Toward a better understanding of the functioning of school systems for better decision-making: The case of primary schools in Togo. In D.W. Chapman & L.O. Mählck (Eds.), *From data to action: Information systems in educational planning* (pp. 147-174). Paris: UNESCO: International Institute for Educational Planning.
- Kulpoo, D., & Coustère, P. (1999). Developing national capacities for assessment and monitoring through effective partnerships. In *Partnerships for capacity building and quality improvements in education. Papers from the ADEA 1997 Biennial Meeting, Dakar, Senegal* (pp. 131-138). Paris: Association for the Development of Education in Africa.
- Kyalo, F.K. (1997). The use of examination results for monitoring performance of schools, districts and provinces. In U. Bude & K. Lewin (Eds.), *Improving test design. Vol. 1. Constructing test instruments, analysing results and improving assessment quality in primary schools in Africa* (pp. 138-144). Bonn: Education, Science and Documentation Centre.
- Le Mahieu, P.G. (1984). The effects on achievement and instructional content of a programme of student monitoring through frequent testing. *Educational Evaluation and Policy Analysis*, 6, 175-187.
- Le Mahieu, P.G., & Leinhardt, G. (1985). Overlap : Influencing what's taught. A process model of teachers' content selection. *Journal of Classroom Interaction*, 21(1), 2-11.
- Lewin, K., & Dunne, M. (2000). Policy and practice in assessment in anglophone Africa: Does globalisation explain convergence? *Assessment in Education*, 7, 379-399.

- Linn, R.L. (1983). Testing and instruction: Links and distinctions. *Journal of Educational Measurement*, 20, 179-189.
- Linn, R.L. (2000). Assessments and accountability. *Educational Research*, 29(2), 4-16.
- Little, A. (1982). The role of examinations in the promotion of the 'Paper Qualification Syndrome'. In *Paper Qualifications Syndrome (PQS) and unemployment of school leavers. A comparative sub-regional study*. Addis Ababa: International Labour Office.
- London, N.A. (1997). A national strategy for systems-wide curriculum improvement in Trinidad and Tobago. In D.W. Chapman, L.O. Mählck, & A.E.M. Smulders (Eds.), *From planning to action: Government initiatives for improving school-level practice* (pp. 133-146). Paris: International Institute for Educational Planning/Oxford: Pergamon.
- Machingaidze, T., Pfukani, P., & Shumba, S. (n.d.) *The quality of education: Some policy suggestions based on a survey of schools, Zimbabwe*. Paris: International Institute for Educational Planning/Harare: Ministry of Education and Culture.
- Madaus, G.F., & Greaney, V. (1985). The Irish experience in competency testing: Implications for American education. *American Journal of Education*, 93, 268-294.
- Madaus, G.F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P.W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 119-154). New York: Macmillan.
- Marope, P.T.M. (1999). Capacity development through ADEA working groups: Applicable practices and lessons. In *Partnerships for capacity building and quality improvements in education. Papers from the ADEA 1997 Biennial Meeting, Dakar, Senegal* (pp. 139-174). Paris: Association for the Development of Education in Africa.
- Mazibuko, E.Z., & Magagula, R. (2003). *Quality of education in Sub-Saharan Africa: A literature review with specific reference to Swaziland c 1991-2001*. A review financed by ADEA through the Educational Network in Eastern and Southern Africa (ERNESA).
- Mioko, S. (1998). *Gender vs socioeconomic status and school location differences in grade 6 reading literacy in five African countries*. Harare: Ministry of Education and Culture.
- Mislevy, R.J. (1995). What can we learn from international assessments? *Educational Evaluation and Policy Analysis*, 17, 419-437.
- Namibia. Ministry of Education and Culture/Florida State University/Harvard University. (1994). *How much do Namibia's children learn in school? Findings from the 1992 National Learner Baseline Assessment*. Windhoek: Ministry of Education and Culture.
- Naumann, J., & Wolf, P. (2001). The performance of African primary education systems: Critique and new analysis of PASEC data for Senegal. *Prospects*, 31, 373-391.
- Ogundare, S. (1988). Curriculum development: A description of the development of the national curriculum for primary social studies in Nigeria. *Educational Studies*, 14, 43-50.
- Omolewa, M., & Kellaghan, T. (2003). Educational evaluation in Africa. In T. Kellaghan & D.L. Stufflebeam (Eds.), *Handbook of educational evaluation* (pp. 465-481). Dordrecht: Kluwer Academic.
- O-saki, K.M., & Agu, A.O. (2002). A study of classroom interaction in primary schools in the United Republic of Tanzania. *Prospects*, 32, 103-116.
- Ouane, A. (Ed.) (2003). *Towards a multilingual culture of education*. Paris: UNESCO: International Institute for Educational Planning.
- Oxenham, J. (1983). What examinations test and emphasize: An analysis of the Lesotho Primary and Junior Secondary Examinations. In *The education sector survey. Annexes to the report of the Task Force*. Maseru: Government of Lesotho.

- Pennycuik, D. (1990a). Factors influencing the introduction of continuous assessment systems in developing countries. In D. Layton (Ed.), *Innovations in science and technology education, Vol. III* (pp. 139-152). Paris: UNESCO.
- Pennycuik, D. (1990b). The introduction of continuous assessment systems at secondary level in developing countries. In P. Broadfoot, R. Murphy, & H. Torrance (Eds.), *Changing educational assessment. International perspectives and trends* (pp. 106-110). London: Routledge.
- Popham, W.J. (1983). Measurement as an instructional catalyst. *New Directions for Testing and Measurement*, 17, 19-30.
- Popham, W.J. (1987). The merits of measurement driven instruction. *Phi Delta Kappan*, 68, 679-682.
- Postelthwaite, T.N., & Wiley, D.E. (1991). *Science achievement in twenty-three countries*. Oxford: Pergamon.
- Rharade, A. (1997). Educational reform in Kenya. *Prospects*, 27, 163-179.
- Rollnick, M., Manyatsi, S., Lubben, F., & Bradley, J. (1998). A model for studying gaps in education: A Swaziland case study in the learning of science. *International Journal of Educational Development*, 18, 453-465.
- Romberg, T.A. (1999). The impact of international comparisons on national policy. In G. Kaiser, E. Luna, & I. Huntley (Eds.), *International comparisons in mathematics education* (pp. 189-199). London: Falmer.
- Ross, K.N. et al. (2000). *Translating educational assessment findings into educational policy and reform measures: Lessons from the SACMEQ initiative in Africa*. Paris: UNESCO.
- Samoff, J. (1999). Cooperation, but limited control and little ownership. In *Partnerships for capacity building and quality improvements in education. Papers from the ADEA biennial meeting, Dakar, Senegal* (pp. 175-204). Paris: Association for the Development of Education in Africa.
- Schiefelbein, E. (1993). The use of national assessments to improve primary education in Chile. In D.W. Chapman & L.O. Mählck (Eds.), *From data to action: Information systems in educational planning* (pp. 117-146). Paris: UNESCO: International Institute for Educational Planning/Oxford: Pergamon.
- Snyder, C.W., Prince, B. Lohanson, G., Odaet, C., Jaji, L., & Beatty, M. (1997). *Exam fervor and fever: Case studies of the influence of primary leaving examinations on Uganda classrooms, teachers, and pupils*. Washington DC: Academy for Educational Development.
- Somerset, A. (1987). *Examination reform in Kenya*. Washington, DC: World Bank.
- Somerset, A. (1988). Examinations as an instrument to improve pedagogy. In S.P. Heynemann & I. Fägerlind (Eds.), *University examinations and standardized testing. Principles, experience, and policy options* (pp. 169-194). Washington DC: World Bank.
- Somerset, A. (1996). Examinations and educational quality. In A. Little & A. Wolf (Eds.), *Assessment in transition. Learning, monitoring and selection in international perspective* (pp. 263-284). Oxford: Pergamon.
- Takala, T. (1998). Making educational policy under influence of external assistance and national politics. A comparative analysis of the education sector policy documents of Ethiopia, Mozambique, Namibia and Zambia. *International Journal of Educational Development*, 18, 319-335.
- Torrance, H. (1997). Assessment, accountability, and standards: Using assessment to control the reform of schooling. In A.H. Halsey, H. Lauder, P. Brown & A.S. Wells (Eds.), *Education: Culture, economy, and society* (pp. 320-331). Oxford: University Press.
- UNESCO. (1990). *World Declaration on Education for All. Meeting basic learning needs*. New York: Author.
- UNESCO. (2000a). *The Dakar Framework for Action. Education for All: Meeting our collective commitments*. Paris: Author.

- UNESCO. (2000b). *Education for All. Status and trends 2000. Assessing learning achievement (2000)*. Paris: Author.
- UNESCO. (2002). *Education for All. Is the world on track? EFA global monitoring report*. Paris: UNESCO.
- UNESCO. (2003a). *Monitoring Learning Achievement (MLA) Project. Update*. Paris: Author.
- UNESCO. (2003b). Southern African Consortium for Monitoring Educational Quality (SACMEQ). Harare: Author.
- V d Merwe, I.F.J. (1999). *Case study on the establishment of a national examinations and assessment system for school examinations in Namibia*. Windhoek: Ministry of Basic Education and Culture.
- Verspoor, A. (1989). *Pathways to change. Improving the quality of education in developing countries*. Washington DC: World Bank.
- Verspoor, A. (1992). *Challenges to the planning of education*. Washington DC: World Bank.
- Wasanga, P.M. (1997). Testing and monitoring procedures developed for primary schools. In U. Bude & K. Lewin (Eds.), *Improving test design. Vol 1. Constructing test instruments, analysing results and improving assessment quality in primary schools in Africa* (pp. 145-157). Bonn: Education, Science and Documentation Centre.
- World Bank. (1988). *Education in Sub-Saharan Africa. Policies for adjustment, revitalization, and expansion*. Washington DC: Author.
- World Declaration on Education for All*. (1990). Adopted by the World Conference on Education for All, Meeting Basic Learning Needs. New York: UNDP/UNESCO/UNICEF/World Bank.