

Table des matières

Présentation du PASEC 5

Chapitre 1 : Ce guide est-il adapté à vos objectifs ? 9

- A. La place de l'évaluation dans le processus de décision et de pilotage des systèmes éducatifs 9
- B. Du problème posé au choix du type d'évaluation 10
- C. Un type particulier d'évaluation : le recueil de données sur un échantillon d'écoles et leur traitement par des analyses multivariées 12

Chapitre 2 : L'exemple des enquêtes PASEC/CONFEMEN sur les facteurs de performance à l'école primaire 15

- A. La modélisation des apprentissages des élèves: 17
- B. Le recueil de données 25
- C. Du protocole d'enquête au montage logistique, financier et institutionnel 27

Chapitre 3 : La conception des instruments d'enquête 37

- A. Les tests d'acquisition 37
- B. Les questionnaires 69

Chapitre 4 : La constitution de l'échantillon 77

- A. Introduction 77
- B. L'échantillon aléatoire simple 78
- C. La stratification 85
- D. L'échantillonnage en grappes 96
- E. Récapitulatif des étapes à suivre pour un échantillonnage en grappes 108

Chapitre 5 : La collecte et la saisie des données 113

- A. Les campagnes de collecte de données 113
- B. La correction des tests 121
- C. La saisie des données 124

Chapitre 6 : L'organisation des données pour l'analyse 151

- A. L'organisation des données dans une seule base 151
- B. Importation et manipulation des données dans un logiciel statistique
..... 163

Chapitre 7 : L'analyse des données 173

- A. La préparation des données 175
- B. L'analyse descriptive des données 181
- C. L'analyse causale 193
- D. Résumé de la méthode d'analyse des données proposée 211

Chapitre 8 : Approche qualitative 219

- A. Pourquoi préconiser une double approche dans l'analyse des systèmes éducatifs ? 219
- B. Elaborer un outil de recueil de données pertinent 226
- C. L'entretien 229
- D. Validation de l'outil de recueil de données 244
- E. Condenser, présenter des données, formuler des conclusions et les communiquer de façon pertinente 247

Chapitre 9 : la rédaction du rapport et la diffusion des résultats .. 255

- A. Le choix d'une équipe et la diversité des points de vue 255
- B. Le rapport scientifique 256
- C. Le séminaire de présentation et de discussion des résultats 259
- D. La diffusion des résultats 260

Chapitre 10 : Variations sur la même méthode 263

- A. Suivis de cohorte 263
- B. Evaluations thématiques 270

Références bibliographiques et contacts utiles 275

Présentation du PASEC

Programme d'Analyse des Systèmes Educatifs des pays de la CONFEMEN

Présentation générale

La plupart des pays d'Afrique Francophone ont connu la même forme de colonisation et par ricochet les mêmes systèmes éducatifs hérités de l'ancienne métropole. Au lendemain des indépendances, chaque Etat, au gré des réformes et des réajustements de programmes, a essayé d'améliorer son système éducatif en tenant compte des exigences nationales.

Les efforts considérables entrepris durant les années 60 et 70 ont permis d'augmenter de manière soutenue le taux de scolarisation. Cependant, au cours des années 80, la conjonction de difficultés économiques et structurelles a abouti à un certain nombre de blocages qui ont remis en cause le mode de développement des systèmes.

La Conférence mondiale sur l'éducation pour tous de Jom-Tien, en 1990, a permis de réaffirmer la priorité accordée à la scolarisation universelle, tout en proclamant l'urgence d'un réexamen des missions, des formes, et des stratégies de cette scolarisation.

En 1990, la 42^{ème} Conférence Ministérielle de la CONFEMEN, tenue à Bamako, a constitué une réponse concrète du monde francophone à ce défi, en mettant un accent particulier sur le caractère indissociable des dimensions quantitatives et qualitatives dans le développement de l'éducation pour tous.

Les Ministres ont alors décidé d'entreprendre en commun un Programme d'évaluation pour aider à la réflexion et à leur action : c'est ainsi que le Programme PASEC a vu le jour, lors de la 43^{ème} Conférence Ministérielle de la CONFEMEN à Djibouti en 1991

Ses objectifs sont les suivants :

- 1) Identifier des modèles d'écoles efficaces et peu coûteux par la comparaison nationale et internationale ;

- 2) Développer dans chacun des états participants une capacité interne et permanente d'évaluation ;
- 3) Diffuser largement méthodes, instruments et résultats en matière d'évaluation.

Concrètement, le Programme PASEC inclut :

1. des opérations d'évaluation (enquête par échantillonnage dans les écoles) ;
2. des formations pour les équipes nationales chargées des évaluations ;
3. un réseau d'information et de suivi géré par le Secrétariat Technique Permanent de la CONFEMEN à Dakar

Les pays participants

Le PASEC comprend à la fois des opérations d'intérêt général (comme la mise à disposition d'outils et de méthodes pour l'évaluation), et des actions d'évaluation menées sur une base nationale, mais coordonnées entre les pays participants.

Ces pays ont été choisis par le Comité Scientifique du PASEC, après appel de candidatures, au fur et à mesure des quatre phases successives de financement qui ont été débloquées depuis la décision de la 43^{ème} session ministérielle de la CONFEMEN à Djibouti en 1991. Le tableau ci-dessous montre la chronologie des opérations en fonction des pays et des phases de financement.

Tableau récapitulatif de l'exécution des PASEC I, II, et III, IV et V

	Phase	Opérations de terrain	Remise du Rapport
Congo	PASEC I	1993/1994	Septembre 1995
Djibouti	PASEC I	1993/1994	Septembre 1995
Mali	PASEC I	1994/1995	Juin 1996
Centrafrique	PASEC II	1994/1995	Mars 1997
Sénégal	PASEC II/IV	1995/1996/1997/1998/1999/2000	
Burkina Faso	PASEC III/IV/V	1995/1996/1997/1998	Octobre 1998
Cameroun	PASEC III	1995/1996	Novembre 1998
Côte d'Ivoire	PASEC III/IV/V	1995/1996/1997/1998/1999/2000	
Madagascar	PASEC IV	1997/1998	Décembre 1999
Guinée	PASEC V	1999/2000	Janvier 2001

Une évaluation du système éducatif togolais devrait également être réalisée dans le cadre du PASEC V au cours de l'année scolaire 2000/2001.

En élargissant le champ d'investigation à plusieurs pays, une plus grande variété de situations scolaires, et donc de solutions, sont évaluées, en même temps que, pour des raisons statistiques, la robustesse des conclusions est améliorée. Ceci nécessite toutefois un bon degré d'harmonisation des instruments (tests et questionnaires) et des protocoles d'enquêtes, surtout en ce qui concerne l'échantillonnage des écoles, les consignes d'administration des tests, et les procédures de compilation et d'analyse des données¹.

L'autre avantage d'une coordination tient aux opportunités de transferts de compétences, via les opérations de formation ou les actions communes à différentes équipes nationales. Ceci favorise, dans chaque pays participant, la construction d'une expertise, et la constitution d'une banque d'instruments et de résultats de référence. L'objectif est dans tous les cas que chacun dispose d'une compétence nationale renforcée pour jauger et juger de la santé et des moyens d'amélioration de son système éducatif.

La plate-forme de contacts ainsi créée entre experts des différents pays constitue d'ores et déjà un réseau d'évaluateurs au service de la qualité de l'éducation dans les pays de la CONFEMEN.

Le présent manuel reprend les étapes principales des évaluations «standard» menées par le PASEC en proposant des solutions aux problèmes concrets posés par la réalisation de telles études. Les points suivants sont ainsi abordés :

- l'élaboration d'une stratégie d'évaluation (chapitres 1 et 2);
- la conception des instruments d'enquête, le choix des écoles ciblées et du protocole d'enquête (chapitres 3 et 4);
- la collecte, la saisie et l'organisation des données recueillies, en prenant pour exemple le système mis en place par le PASEC (chapitres 5 et 6);
- diverses voies possibles d'analyse des données recueillies (chapitre 7) et de diffusion des résultats (chapitre 9).

Enfin, sont proposés des méthodes complémentaires, dont la mise en oeuvre est actuellement en cours au PASEC : l'approche qualitative, conçue comme complément à l'enquête statistique (chapitre 8) et des stratégies d'enquête sensiblement différentes (suivi de cohorte, évaluation thématique ; chapitre 10).

¹ Le site Internet de la CONFEMEN propose une présentation plus détaillée du PASEC et des résultats obtenus, ainsi que le téléchargement des instruments utilisés, à l'adresse suivante :

Chapitre 1 : Ce guide est-il adapté à vos objectifs ?

Ce guide a été écrit à l'intention de deux types de publics :

- d'une part les responsables d'évaluation, afin de leur fournir des outils d'aide à la réalisation d'une évaluation des facteurs de performance à l'école primaire
- d'autre part les acteurs et usagers de l'école, désireux de mieux savoir comment sont produits les résultats des chercheurs

A. La place de l'évaluation dans le processus de décision et de pilotage des systèmes éducatifs

L'évaluation constitue une aide pour les décideurs. En matière d'éducation primaire, elle a surtout pour but de donner des indications sur les performances des élèves, sur les inégalités dans les conditions de scolarisation et dans les résultats, et sur les facteurs les plus susceptibles d'améliorer les apprentissages.

De plus en plus, les responsables éducatifs s'efforcent de mieux intégrer l'évaluation au processus de décision : pour mieux cerner les besoins, pour choisir les mesures les plus efficaces, pour arbitrer entre objectifs concurrents, comme par exemple le nombre d'élèves scolarisés et la qualité de l'éducation.

L'évaluation est particulièrement pertinente quand des innovations sont proposées. Elle sert alors à concevoir des expérimentations, et le cas échéant à valider leur réussite.

L'évaluation est un élément important du débat démocratique, en présentant aux partenaires de l'éducation les faits et les tendances de la manière la plus objective possible.

Pour tenir sa place dans tous ses usages et objectifs, l'évaluation doit rester fidèle à son principe premier, qui est de fonder toutes ses conclusions sur des faits et des données d'observation recueillis sur le terrain.

B. Du problème posé au choix du type d'évaluation

Beaucoup de pratiques différentes sont regroupées sous le terme d'évaluation. Le choix d'une méthode particulière dépend du problème posé, des moyens et des compétences disponibles.

Au niveau d'une classe, si le problème est de tester le niveau des élèves, ou de prendre certaines décisions (sur le rythme d'apprentissage, sur les redoublements,...), l'évaluation peut être le fait du maître qui interroge ses élèves, à l'oral ou à l'écrit.

Au niveau d'un établissement, si l'objectif est de vérifier la qualité de l'enseignement, d'améliorer les pratiques pédagogiques, ou de noter les personnels d'éducation, l'évaluation est souvent réalisée par un inspecteur ou un conseiller pédagogique. Celui-ci assiste à des classes, réalise des entretiens avec les maîtres, avec le directeur, puis produit un rapport.

Au niveau de l'administration centrale, si l'objectif est d'améliorer l'organisation administrative générale du ministère, la gestion des personnels enseignants, ou l'utilisation des fonds d'investissement, un audit peut être conduit, soit en interne (par des agents du ministère), soit en externe (par exemple par un cabinet d'experts).

Toutes ces pratiques peuvent légitimement se réclamer de l'évaluation.

Certaines questions ne peuvent pourtant être traitées de manière approfondie que par le passage à un système de collecte d'information plus systématique.

Il en va en particulier des questions qui mettent en jeu les rapports entre la situation de scolarisation et les performances scolaires des élèves. Par exemple “ le redoublement est-il efficace ? ”, ou “ la politique de recrutement de contractuels a-t-elle un effet sur le niveau des élèves ? ”, ou “ dans quelle mesure le système des classes à double flux (ou les classes à grands effectifs, ou le bilinguisme) handicape-t-il l'apprentissage ? ”, ou bien encore “Existe-t-il un handicap spécifique des filles par rapport aux garçons, et pourquoi ? ”.

Les acteurs du système (maîtres, inspecteurs, pédagogues) peuvent déjà donner beaucoup d'éléments de réponse à toutes ces questions. Un enquêteur extérieur au système peut également, au moyen de visites et d'entretiens, apporter un éclairage spécifique. Les méthodes qualitatives des sciences sociales permettent d'ordonner et de systématiser de telles analyses. Malgré tout, d'un chercheur à un autre, il pourra exister des différences de conclusions assez fortes, car la méthode laisse, et c'est heureux, une large part à l'appréciation et au jugement. Une argumentation, même informée, peut toujours être contestée par une autre, sans véritable critère accepté par tous permettant de trancher entre les deux.

En revanche, une approche statistique permet d'apporter des éléments plus difficilement contestables, même s'ils sont moins fins. Les choix et les interprétations sont toujours présents, mais ils sont mieux isolés du processus de collecte et de traitement des données. La contestation des conclusions a toujours droit de cité, mais, la plupart du temps, elle laisse intact le constat statistique qui se dégage, et dont tous doivent rendre compte.

Dans ce cas, la démarche passe toujours par la collecte de données à l'aide d'instruments d'enquête (tests de niveau, questionnaires), distribués à un échantillon aléatoire ou raisonné d'élèves, de maîtres, de directeurs, etc .

Les données recueillies sont ensuite traitées de manière systématique, et il en ressort des conclusions soit descriptives (comme le pourcentage d'élèves qui n'ont pas atteint le niveau jugé minimal par les responsables pédagogiques), soit prédictives (comme l'influence de telle ou telle variable, comme la taille de la classe, sur les performances des élèves).

Il est bien entendu que chaque évaluation a recours à des outils d'analyse spécifiques, et il appartient au commanditaire de l'étude de préciser s'il souhaite que l'une ou l'autre approche soit privilégiée. Il ne s'agit pas pour nous d'opposer ces différentes méthodes. En particulier, l'opposition entre approches qualitative et quantitative paraît stérile, puisque celles-ci ne peuvent se passer l'une de l'autre. Comment expliquer des résultats quantitatifs en ignorant le fonctionnement qualitatif du système scolaire? Et comment s'assurer que ce qui a été mis à jour sur une petite échelle par l'approche qualitative est généralisable à plus grande échelle, sinon en adoptant, peu ou prou, l'attitude du statisticien? Les évaluations en cours de la CONFEMEN, en développant

d'avantage leur partie qualitative, s'engagent significativement en faveur de cette complémentarité entre approches qualitative et quantitative.

C. Un type particulier d'évaluation : le recueil de données sur un échantillon d'écoles et leur traitement par des analyses multivariées

Il est possible d'établir une correspondance générale entre des degrés successifs d'objectifs assignés à l'évaluation, et des degrés successifs de complexité des techniques statistiques adaptées.

Ainsi, au fur et à mesure que l'objectif de l'évaluation passe du diagnostic à la proposition, il est nécessaire de passer de l'analyse univariée (idéale pour établir l'état du système), à l'analyse bi-variée (plus spécialement destinée à repérer les inégalités), puis à l'analyse multivariée (seule efficace pour déterminer les causes, et donc pour opérer des changements).

Les paragraphes qui suivent ont pour but de convaincre le lecteur que l'analyse univariée, indispensable, a tout intérêt à être complétée par une analyse bi-variée (mettant en relation deux grandeurs ou qualités, comme le score d'un élève en fonction de son sexe), puis, plus largement, multivariée (mettant en relation plus de deux grandeurs ou qualités : comme le score, le sexe, l'âge, l'origine sociale,...).

Prenons le cas d'un ministre désireux de mesurer le niveau moyen des élèves de fin de cycle primaire, en choisissant pour cela un test de référence. La solution lourde consiste à faire passer à tous les élèves du pays et du niveau considéré un même test. Cela n'est pas vraiment envisageable. Il est plus efficace de ne retenir qu'un nombre restreint d'élèves choisis par tirage au sort. Si les règles sont respectées, c'est-à-dire si chaque élève du pays a la même probabilité de faire partie de l'échantillon, le niveau moyen peut être estimé avec une grande précision.

Estimer une grandeur moyenne, comme un niveau de performance des élèves, relève de l'analyse univariée. C'est une démarche utile pour établir un diagnostic, et donc, éventuellement, la nécessité d'une intervention.

Pour aller plus loin, les évaluations font en général une large place à des résultats “ bi-variés ”. Par exemple, au-delà de la mesure univariée du score moyen des élèves, l'analyse fournira une mesure bi-variée du score selon le contexte (score moyen des élèves des écoles urbaines, score moyen des élèves des écoles rurales).

Les analyses bi-variées apportent une précision supplémentaire par rapport aux analyses univariées. Dans le cas envisagé plus haut, la mesure univariée du niveau moyen des élèves donne une indication sur l'effort à accomplir pour atteindre les objectifs du programme, mais l'analyse bi-variée du niveau selon le milieu permet d'aller plus loin, et de déterminer si les élèves des écoles en milieu rural ont davantage de lacunes que leurs camarades des villes. Nous sommes déjà un peu plus proches des mesures à prendre que dans le cas de l'analyse univariée.

Néanmoins, pour fournir vraiment une aide à la décision, il faut passer à une analyse multivariée (qui met en relation plus de deux variables), seule susceptible d'indiquer les relations de cause à effet.

Reprenons un exemple pour illustrer cette nécessité de passer à l'analyse multivariée dès qu'il s'agit d'estimer l'impact d'une variable ou d'une mesure. Admettons en l'occurrence l'existence d'un résultat bi-varié selon lequel les élèves qui progressent le plus sont aussi ceux dont les maîtres sont les plus diplômés. On serait alors tenté de proposer une politique de recrutement des maîtres exigeante sur le diplôme pour élever le niveau.

Le risque de se tromper d'effet, dans ce cas, est important. En effet, le lien statistique (la corrélation, en termes techniques) entre deux variables n'est pas une preuve de causalité : il est trop tôt, à ce stade, pour conclure qu'un niveau élevé de diplôme du maître est directement lié à de meilleurs résultats des élèves. Dans ce cas, on découvrira peut-être, en incluant la variable de milieu dans notre analyse, que la plupart des maîtres qui ont un bon niveau de diplôme ont été nommés dans les villes, et qu'ils ont affaire à des élèves plus avantagés. Ils n'ont donc pas de mal à obtenir de bons résultats de ces élèves, à la différence de leurs collègues de milieu rural, confrontés à un public d'élèves moins avantagés et à des conditions de travail plus difficiles.

Si l'analyse ne prend pas en compte les effets de contexte, la conclusion concernant l'effet de telle ou telle variable sur telle ou telle autre risque d'être

fausse. Pour prendre en compte cet effet, il faudra passer de l'analyse bi-variée à l'analyse multivariée ; dans la seconde, il est possible de dégager l'effet propre de chaque variable (le milieu de l'élève, le niveau de diplôme du maître) sur la variable que l'on cherche à expliquer, compte tenu de tous les effets de contexte qui jouent en même temps. C'est cette prise en compte des effets de contexte qui justifie qu'on qualifie parfois l'analyse multivariée d'analyse " toutes choses égales par ailleurs "

En définitive, l'analyse multivariée est beaucoup plus indiquée que l'analyse bi-variée pour fournir au décideur des indications fiables sur les mesures les plus efficaces.

Le tableau qui suit propose un résumé des apports particuliers des analyses univariées, bi-variées et multivariées.

	Analyse univariée	Analyse bi-variée	Analyse multivariée
Usage privilégié	Diagnostic global	Diagnostic affiné	Choix d'une action efficace
Exemple de question	Les objectifs de programme sont-ils atteints ?	Les élèves des classes rurales ont-ils un niveau plus faible ?	Est-ce que des maîtres plus diplômés permettent d'améliorer les apprentissages des élèves ?
Nombre de variables prises en compte simultanément dans l'analyse	1	2	3 ou plus
Exemple de variables	Score	Score ; Contexte (rural ou urbain)	Score ; Contexte ; Diplôme du maître ; (...)

L'ambition de ce guide est d'amener ses lecteurs aux techniques de réalisation d'une analyse multivariée.

Ce choix se justifie par la rareté du nombre des évaluations orientées vers la prise de décision, alors que les évaluations orientées vers le diagnostic sont beaucoup plus fréquentes.

Ce choix d'objectif ne vise pas à minimiser les problèmes d'échantillonnage et d'estimation dans le cas des analyses univariées et bi-variées, mais à amener au plus vite, et sans complexité inutile, à un usage accru de l'analyse multivariée en éducation.