

LES EFFETS THEORIQUES ET REELS DE L'EVALUATION STANDARDISEE

Nathalie Mons

CHAMP

Évaluation des compétences des élèves et de l'efficacité du système éducatif

Pilotage de l'institution

Indicateurs

redevabilité

Nouvelle politique de management

AVERTISSEMENT:

La lecture de cet article requiert de la part du lecteur qu'il ait bien à l'esprit que la culture de l'évaluation standardisée s'est développée aux EU et dans les pays anglo-saxons où les programmes, quand ils existent, ne sont pas uniformisés au niveau national. D'autre part, il est à noter qu'une certaine forme de pensée plutôt libérale colore la politique anglo-saxonne s'appuyant aussi sur un courant de pensée économique très puissant conditionnant les outils qui lui sont associés et les mesures qui en découlent.

INTRO

«D'un simple outil neutre d'évaluation des compétences et des savoirs des élèves, il s'est imposé comme un instrument majeur d'une nouvelle régulation politique des systèmes éducatifs... Il est devenu un outil de pilotage», p5.

Évaluation standardisée (ES= évaluation qui a fait l'objet d'un traitement statistique qui rend relativement fiables les résultats dans la mesure où les items sont censés réellement mesurer ce qu'ils mesurent cad que les résultats qui en découlent se répartissent, comme la majeure partie des caractéristiques humaines, en courbe de Gauss) doit être mise en relation avec 4 facteurs:

1. La centration sur l'aspect quantitatif des apprentissages et priorité donnée aux des obj. cognitifs au détriment des obj de socialisation.
2. Importance grandissante du contrôle des enseignants par les personnels administratifs
3. Évolution de la répartition des pouvoirs entre acteurs locaux et décideurs
4. Développement de la redevabilité de l'école (*réf: Accountability, npp (New public management...)*)

Cette nouvelle posture, ces nouveaux dispositifs ont à être questionnés au vu de la place qu'ils prennent de plus en plus et des débats vifs qu'ils suscitent.

D'où la réalisation de ce rapport.

Quels objectifs? Quelle efficacité? Quel coût? Quelle pertinence? Quels impacts sur la population scolaire?

I. LES EFFETS ATTENDUS: CADRE THEORIQUE POLITIQUE ET PEDAGOGIQUE DE L'EVALUATION STANDARDISEE

L'évaluation standardisée a été généralisée dans les pays développés, «elle a acquis un nouveau statut politique qui fait d'elle un instrument central dans la régulation des systèmes»,p7.

1. Le cadre théorique politique de l'évaluation standardisée: un instrument de pilotage des systèmes éducatifs

1.1. Évaluation standardisée et rénovation de la gestion de l'État: npm et le mouvement de policy évaluation

Aux EU, dans un contexte de crise économique qui remet en cause l'action de l'État et la gestion des fonds publics, émerge une nouvelle politique appelée New Public Management (npm) qui prône l'assainissement des finances publiques à partir de thèses dites monétaristes, les évaluations standardisées émergent et colorent les politiques mises en place aussi dans le secteur de l'Éducation. Il s'agit bien de mesurer l'efficacité de l'Institution pour justifier des coûts engagés.

C'est une posture très rationnelle qui envisage les institutions scolaires dans une logique telle que suit, héritée d'une démarche et forme de pensée très proches de l'explicative causale des sciences expérimentales:

- . Productions service public sont mesurables
- . Elles peuvent être mesurées par des outils scientifiquement validés
- . Les acteurs publics sont redevables devant leurs responsables et les

citoyens de leur action.

- . Evaluation sur les résultats et non seulement sur les ressources.
- De sorte que l'ES permettrait ;
- . Évaluation des acquis des E en référence à des objectifs cognitifs
 - . Lien entre les acteurs du système et les responsables en charge de la conception.
- les parents.
- . Médiatisation de l'information vers l'extérieur du système entre autres

1.2. La logique économiste de l'ES et le courant pragmatique de la school effectiveness

Dans perspective de la théorie du capital humain, l'investissement économique fait dans l'éducation est un investissement pour l'avenir dans la mesure où plus un pays est composé d'individus diplômés plus les richesses nationales sont importantes. Ainsi il faut s'interroger sur la réalité des apprentissages et la valorisation qui est faite des deniers investis dans l'éducation. Il faut aussi rendre les individus responsables de leur action professionnelle.

Il est fait abstraction dans ces mesures, des éléments afférents au contexte qui peuvent être des éléments d'analyse pertinents, plus fins et plus représentatifs de la réalité.

L'ES est censée influencer les pratiques d'enseignement et les faire évoluer dans le bon sens.

L'outil de mesure, ES, est donc, dans différentes perspectives:

- . un outil d'information interne et externe à l'institution
- . un outil qui décide des carrières scolaires des élèves et leur orientation
- . un outil de mise en cohérence des actions à tous les niveaux
- . un outil de direction cognitive et de contrôle des agents
- . un outil de transfert de responsabilité

On se trouve dans un contexte de libéralisation de l'École.

2. Le cadre théorique pédagogique de l'évaluation standardisée : des modèles encore en construction

« Beaucoup d'auteurs pointent du doigt le fait que les théories expliquent rarement précisément par quels processus intermédiaires l'ES est censée conduire à une progression des acquis des élèves. , p11 »

Le modèle de l'accountability (signifie la logique du rendre compte) **dure** implique résultats/récompense ou sanction pouvant aller jusqu'à la fermeture de l'établissement.

Il est fondé sur les mécanismes suivants:

- . Les élèves doivent fournir un travail plus important car passage en classe supérieure ou certification dépendent du résultat au test.
- . Les E sont responsabilisés et peuvent donc travailler davantage à la réussite de leurs élèves et améliorer leur professionnalisme grâce à des formations, à des échanges et à des analyses réflexives.
- . les ES réduisent les inégalités scolaires en imposant des standards communs à tous les élèves quels que soient les contextes.«Confrontées à des résultats considérés comme faibles mais à des objectifs identiques, les équipes pédagogiques des écoles accueillant des publics difficiles se trouveraient dynamiser par ces nouveaux défis collectifs », p12.

Commentaires : On se trouve dans une perspective béhavioriste dure qui pourtant dans la théorie de la motivation a été, de manière très large, mise à mal car les effets obtenus ne sont pas pérennes ni performants voire entraînent des effets pervers. Aujourd'hui on s'oriente plutôt vers des modèles d'influence plus doux. Ils prennent en compte la part des individus en les responsabilisant, en les renseignant sur la qualité de leur action , en leur proposant des buts de tâche...mettre en contradiction motivation extrinsèque et motivation intrinsèque pour décrire comment les modèles peuvent avoir un effet plutôt qu'un autre.

C'est un **modèle accountability douce**, qui se développe plutôt dans les pays européens.

En Allemagne, on est alors plutôt orienté vers des standards de résultats en terme d'objectifs d'enseignement à atteindre, mesurés par des tests. Ceux ci renseignent les parents, les élèves et les E sur les champs d'apprentissage prioritaires. Ils ne décident en rien des carrières scolaires des élèves et donnent aux E une orientation pour enseigner. Ils sont utilisés comme indicateurs et non comme outil donnant lieu à sanction ou récompense, ni ne font l'objet d'une médiatisation large.

En France, on parle plutôt d'«effet miroir» , cf. Thélot, cad que les évaluations rendent une image des effets de la pratique et non une explication de la réussite scolaire.

Mais que l'on se pose dans un modèle ou l'autre, un certain nombre de présupposés sont présents:

- . Fiabilité de la mesure
- . Inexistence de variable parasite
- . La contrainte que représente le test améliore la performance de l'élève
- . Le test est une aide pour les E
- . Les tests sont des indicateurs fiables pour le pilotage du système
- . les tests et leurs résultats sont lisibles et compréhensibles par tous.

Pourtant un large champ de recherches ne partagent pas ces points de vue.

II. LES EFFETS REELS DE L'EVALUATION STANDARDISEE SUR LES PERFORMANCES DES SYSTEMES EDUCATIFS

C'est aux EU que la controverse et le débat ont été les plus larges, à l'occasion des réformes du district de Chicago et de l'état du Texas. Les recherches montrent « l'absence de consensus empirique sur l'impact des dispositifs ... et un ensemble de chausse trappes méthodologiques... enlevant toute validité scientifique », p 17.

Comme les réussites des élèves aux tests en dixième année ont été bien supérieures entre 1994 et 2002 passant de 28 à 78%, il en a été déduit que les effets du test étaient positifs.

Pourtant, d'autres recherches ont mis en perspectives les résultats des élèves aux tests locaux et nationaux et mis en évidence que leurs résultats aux évaluations nationales étaient soit beaucoup moins bons soit sans différence significative sur la même période.

On en retiendra qu'«une évaluation locale ne peut servir à la fois de moyen d'intervention et être son propre outil d'évaluation», p 17.

D'autre part, les écarts de progrès aux tests locaux sont expliqués par:

- . le phénomène que l'on appelle «teaching to the test»
- . le fait qu'un certain nombre d'élèves faibles avaient été écartés du test
- . les E pouvaient accordés plus de temps d'enseignement aux élèves les plus

susceptibles de progresser de manière significative soit les élèves moyens laissant de côté particulièrement les plus faibles.

« Au total, les relations entre efficacité et dispositifs de testing n'apparaissent ni uniformément significatives ni automatiques », p 21. De plus, si l'analyse statistique prend en compte en isolant la variable contexte économique ou, et démographique ou la variable développement économique alors la liaison accountability / performances scolaires disparaît. De même, il n'existe pas de lien entre testing et inégalités scolaires sauf pour un type particulier: les examens nationaux externes dans le secondaire qui seraient un moyen pour amener tous les élèves vers les mêmes objectifs, « peut être parce qu'elles homogénéisent les exigences académiques qui sont imposées aux E et limitent les dérives curriculaires dans les établissements qui accueillent les élèves des milieux défavorisés. »,p23.

En ce qui concerne l'efficacité de ces politiques d'ES, il est intéressant de s'interroger sur leur pertinence vu les coûts astronomiques qu'elles impliquent.

III. EVALUATION STANDARDISEE ET PROCESSUS EDUCATIFS : LES COMPORTEMENTS DES ENSEIGNANTS, CADRES INTERMEDIAIRES ELEVES ET PARENTS FACE AU TESTING

1. Les Enseignants: une tendance à la résistance à la culture de l'évaluation standardisée quantitative

Les ES peuvent avoir des effets positifs sur l'enseignement quand les E y adhèrent, qu'ils ne sont pas contraignants ni censeurs car ils permettent de mieux organiser les contenus, de mieux comprendre ce qui doit être enseigné. Cela est montré dans des pays comme France, EU, GB et Suède. « La perception générale des standards... et des dispositifs de test semble positive car ils constituent des guides clairs pour mettre en œuvre les programmes, empêchent l'apparition de fortes inégalités, mettent l'accent sur les résultats réels des élèves et favorisent un travail d'équipe », p 27.

Pour que l'ES soit bien perçue par les E, trois conditions (cf Demailly 2001):

- . Faire participer les E à l'élaboration de l'ES
- . Les objectifs sont démocratiques
- . Les promoteurs du projet performants et engagés

Ces conditions n'étant pas réunies, se font jour des effets pervers surtout si les tests sont associés à des forts enjeux conduisant à « une évolution défavorable des pratiques pédagogiques,(...) un sentiment de déprofessionnalisation,(...)qui se traduit par des phénomènes de démotivation des enseignants. », p27. Entraînant aussi, «teaching to the test», des tricheries, stress, un rétrécissement des contenus d'enseignement autres que les matières évaluées et favorisant ainsi l'élaboration de compétences basiques, la mémorisation rapide «au détriment des autres missions de l'école (socialisation, développement de la créativité, autonomie, participation à la vie citoyenne)», p28. Pouvant aller jusqu'à la sélection de publics recrutés par l'école, cela est accentué encore quand les établissements sont placés dans une logique de concurrence.

La démotivation des E semblent prégnante dans les milieux accueillant des élèves défavorisés, pouvant aller jusqu'à « un profond sentiment de déprofessionnalisation. », p30.

2. Les cadres du système éducatif: une prise en main de l'outil

De manière générale, surtout aux EU, en Belgique et en GB, ils sont favorables aux ES; en France, cela est vrai aussi pour le primaire, moins pour le secondaire. Le premier utilisant plus volontiers les résultats des ES que le second.

3. Les élèves et le fardeau des tests

Les tests sont censés améliorer la motivation des élèves au moins en termes de motivation extrinsèque (*dont on connaît les limites !!!*).

Ils semblent générer anxiété. Ils peuvent jouer un rôle négatif dans les carrières scolaires des élèves « obligeant certains à redoubler, stigmatisant la difficulté scolaire...allant même jusqu'à entraîner une augmentation des abandons scolaires », p33. Ces phénomènes seraient amplifiés dans le cas des élèves défavorisés.

4. Les parents: une réception positive du testing mais des attentes pour l'école qui vont au delà

L'outil fournit informations, constitue une base de discussion mais pensent qu'ils ne suffisent pas en soit.

CONCLUSION:

« Il semble que cet instrument, certainement parce qu'il est associé à des pratiques traditionnelles de l'école- l'évaluation et la notation-n'entraîne pas un positionnement -positif ou négatif-à priori. L'outil apparaît dans un premier abord, neutre et semble le demeurer quand les conséquences qui y sont attachées demeurent faibles. », p35. En tout état de cause, le cadre théorique des effets de l'ES sur « les performances des élèves demeure faible, les mécanismes intermédiaires doivent être détaillés et évalués empiriquement », p35. Les recherches empiriques ne dessinent pas un consensus. Les effets de ces instruments paraissent plutôt aléatoires. « L'évaluation des coûts de ces politiques est peu explorée », p35.

C'est pourquoi il faudrait:

Une bonne connaissance des éléments cruciaux qui président à la mise en place et à la conception de ces tests est indispensable pour ceux qui souhaitent utiliser ce type de pratique (cf Mons&Pons, 2006):

- . CONCEPTION,PILOTAGE, CONTEXTUALISATION:

Un questionnement de fond sur la liaison entre réformes et ES, en lien avec les programmes, la formation des enseignants, le financement des écoles et le développement d'un accompagnement dans l'analyse et la régulation qui font suite aux ES.

- . BUT VISE, NATURE DU TEST

Poser la question de l'utilisation du test, quel nombre d'épreuves pour avoir une image fiable de ce que l'on cherche à savoir, développer donc un outil qui s'appuie sur des connaissances et un traitement statistique indépendant et de haut niveau. « Construire une épreuve unique pour atteindre des objectifs multiples (direction des carrières scolaires, pilotage des établissements,

évaluation générale du système scolaire) conduit à des dysfonctionnements. », p36.

. INTEGRATION DES ENSEIGNANTS A LA CONCEPTION, A L'ANALYSE ET A LA
REGULATION DU DISPOSITIF

Afin de développer un outil contextualisé mais aussi de développer la culture de l'évaluation.

. LA QUESTION DE LA MEDIATISATION

Pour qui? Pour quoi?

De manière générale, l'utilisation des ES pose la question des modalités de mise en place et de développement des politiques publiques d'éducation. Entre un évaluation gestionnaire, gérée par l'administration et une évaluation démocratique, gérée par le politique, « un nouveau modèle d'évaluation professionnelle (dispositif d'évaluation formative donnant une large place à l'intervention des professionnels de l'éducation qui en sont les premiers utilisateurs) peut certainement compléter ce tableau traditionnel. », p37.

Notes:

Il semble indispensable pour avoir un regard distancié par rapport au sujet de mettre en perspective ces pratiques plutôt anglo saxonnes -même si elles ont tendance à aussi s'installer chez nous dans une moindre mesure- avec les recherches autour de la motivation et leurs résultats sur les effets du contrôle, de la contrainte, de la récompense et de la sanction qui entraînent le plus souvent démotivation, externalisation des objectifs, effets de courte durée cad disparition de la pratique quand celui qui dispense l'injonction n'est plus là. Enfin, un petit détour par la théorie des minorités actives et celle de l'innovation donnerait peut être, sûrement?, des pistes autrement plus ambitieuses!!