

## DES ÉPREUVES DE RÉFÉRENCE AU SERVICE DE L'ÉVALUATION DES ENSEIGNANTS

Martine Wirthner, collaboratrice scientifique à l'IRDP/Neuchâtel  
Ladislav Ntamakiliro, collaborateur scientifique à l'URSP/Lausanne  
et au Département de l'Instruction publique/Genève

### Introduction

Officiellement, les épreuves externes ne sont pas destinées à l'évaluation des enseignants ni même à l'évaluation des écoles. L'administration scolaire tient volontiers ce discours aux enseignants et à leur représentation syndicale qui ne manquent pas de s'interroger sur l'utilisation possible des informations concernant les performances des élèves à ces épreuves. Ce discours est sincère dans la mesure où, en Suisse romande, les résultats des élèves aux épreuves externes ne sont pas exploités aux fins de l'évaluation certificative des enseignants, encore moins en vue d'établir un palmarès des écoles. En revanche, le rôle des épreuves externes dans l'évaluation formative des enseignants et dans le monitoring des résultats des établissements scolaires est clairement revendiqué et dans une certaine mesure assuré. L'évaluation formative sinon certificative des enseignants et des établissements est par ailleurs plus ou moins explicitement envisagée dans la perspective des épreuves de référence romandes, un projet de la CIIP, comme dans celle des épreuves du projet HarmoS de la CDIP

L'apport et les limites des épreuves externes dans la perspective de l'évaluation des enseignants et des établissements étaient au centre de cet atelier organisé en deux étapes : une communication suivie d'un débat. Dans le texte ci-après nous présentons d'abord l'essentiel du contenu de la communication. Nous rendons compte ensuite, succinctement, des questions principales et des idées maîtresses exprimées au cours des deux séances de l'atelier.

### Aspects conceptuels

A l'opposé des évaluations assurées par les enseignants et dont la portée est interne à la classe, les épreuves externes sont élaborées par l'administration scolaire et concernent le plus souvent tous les élèves du pays ou du canton. En Suisse romande, ces épreuves externes sont diversement qualifiées : épreuves de référence, épreuves cantonales de référence, épreuves cantonales, épreuves communales.

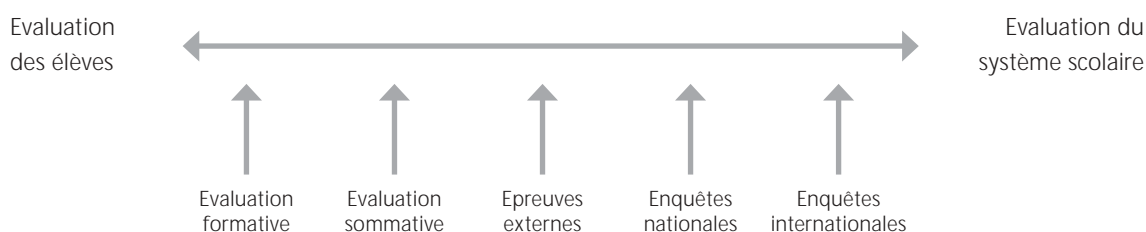
Derrière cette diversité de termes, on distingue deux catégories d'épreuves externes selon qu'elles sont standardisées ou non. D'après De Landsheere (1979), une épreuve est standardisée si tous les élèves du degré concerné sont soumis à l'épreuve et que les conditions d'administration et de correction sont uniformisées.

Les épreuves externes non standardisées consistent essentiellement en banques d'épreuves. Quand ils les utilisent, les enseignants en adaptent les conditions de passation et les critères de correction en fonction de leur enseignement et des caractéristiques de leurs élèves.

Le problème de l'évaluation des enseignants ne concernant guère les épreuves externes non standardisées, nous nous limitons ici aux épreuves externes standardisées.

Les fonctions des épreuves externes standardisées sont multiples. On s'accorde avec Weiss (2003) pour dire que ces épreuves sont conçues et utilisées selon deux logiques distinctes : celle de l'évaluation de l'élève, d'un côté, celle de l'évaluation du système scolaire, de l'autre. La figure 1 représente la position des épreuves externes sur un continuum de pratiques d'évaluation scolaire relevant des deux logiques. D'une part, les épreuves externes standardisées contribuent à l'évaluation des élèves en complément aux évaluations formatives et sommatives assurées par les enseignants. D'autre part, les épreuves externes participent à l'évaluation du système scolaire aux côtés des enquêtes nationales et internationales relatives au rendement scolaire.

*Figure 1 : Position des épreuves externes sur un continuum entre les pratiques d'évaluation des élèves et les pratiques d'évaluation du système scolaire*



D'un canton à l'autre, d'un pays à l'autre, la composante « évaluation des élèves » est plus ou moins importante par rapport à la composante « évaluation du système scolaire », mais la double finalité des épreuves externes est une constante.

Il convient de préciser que c'est dans la mesure où elles contribuent à l'évaluation du système scolaire que les épreuves externes peuvent être mises au service de l'évaluation des enseignants, venant alors compléter les enquêtes nationales ou internationales considérées comme plus appropriées pour remplir cette fonction.

## Situation des pratiques d'évaluation externe en Suisse romande

Le canton de Berne est le seul canton romand qui ne pratique pas d'évaluation externe (nos données concernent au plus tard l'année scolaire 2006/2007). Tous les autres cantons organisent des épreuves externes standardisées en 6<sup>e</sup> et/ou dans d'autres degrés de l'enseignement obligatoire (tableau 1). Le recours aux épreuves externes standardisées en 6<sup>e</sup>, soit au moment où s'opère l'orientation des élèves dans différentes filières du secondaire inférieur est une vieille tradition. L'administration de telles épreuves dans d'autres degrés est en revanche une tendance relativement récente. De plus en plus de cantons mettent en place des épreuves externes standardisées dans les deux degrés marquant la fin des cycles de deux ans d'école primaire, la 2<sup>e</sup> et la 4<sup>e</sup> ainsi que dans le degré précédant la fin de l'école obligatoire, soit la 8<sup>e</sup>. L'évolution récente des pratiques d'évaluation externe dans les cantons de Fribourg, Genève et Vaud illustre bien cette tendance.

*Tableau 1 : Situation des épreuves externes dans l'enseignement obligatoire par degré et par canton en 2006-2007*

Degrés	1	2	3	4	5	6	7	8	9
Berne									
Fribourg		●		●		●			
Genève		●		●		●	●	●	●
Jura						●			
Neuchâtel						●			
Valais				●		●		●	●
Vaud		●		●		●		●	

Dans le canton de Fribourg (tableau 2), en plus du traditionnel test d'orientation passé en 6<sup>e</sup>, deux nouvelles épreuves externes standardisées sont administrées alternativement en 4<sup>e</sup> et en 2<sup>e</sup> depuis l'année scolaire 2005-2006. Celles-ci sont destinées moins à l'évaluation des élèves qu'au suivi de la mise en place des cycles de deux ans ainsi que d'autres innovations pédagogiques.

*Tableau 2 : Evolution de la situation des épreuves externes dans le canton de Fribourg*

Degrés	1	2	3	4	5	6	7	8	9
Avant 2005						●			
En 2005-06				●		●			
En 2006-07		●				●			

Dans le canton de Genève (tableau 3), le traditionnel dispositif d'évaluation externe en 6<sup>e</sup>, 7<sup>e</sup>, 8<sup>e</sup> et 9<sup>e</sup> s'est récemment élargi à la 2<sup>e</sup> puis à la 4<sup>e</sup>, en raison de la rénovation de l'école primaire.

*Tableau 3 : Evolution de la situation des épreuves externes dans le canton de Genève*

Degrés	1	2	3	4	5	6	7	8	9
Avant 2000						●	●	●	●
En 2000-01		●				●	●	●	●
En 2005-06		●		●		●	●	●	●

En 2000, l'évaluation externe est assurée également en fin de 2<sup>e</sup> année, en français et en mathématiques. En 2006, suite aux ajustements introduits dans la réforme de l'école primaire, le dispositif d'évaluation externe s'est encore élargi. Le cycle moyen de quatre ans, jugé trop long, ayant été scindé en deux cycles, de nouvelles épreuves externes ont été introduites en fin de 4<sup>e</sup>, en français, en mathématiques et en allemand.

La situation du canton de Vaud (tableau 4) illustre encore mieux cette tendance vers plus d'épreuves externes. Après une longue période d'utilisation importante d'épreuves externes standardisées aux fins de l'orientation et de la sélection des élèves à l'entrée du secondaire inférieur, ce canton avait un temps rompu avec cette pratique. Le système d'évaluation mis en place dans le cadre de la réforme scolaire de 1997 était essentiellement formatif et qualitatif alors que les épreuves externes standardisées sont sommatives et quantitatives.

*Tableau 4 : Evolution de la situation des épreuves externes dans le canton de Vaud*

Degrés	1	2	3	4	5	6	7	8	9
Avant 2000					●	●			
En 2000-01		●		●		●			
En 2005-06		●		●		●		●	

En 2002, pour pallier les faiblesses du nouveau système d'évaluation, les épreuves externes standardisées ont été réintroduites en 5<sup>e</sup> et en 6<sup>e</sup> année. Le but principal était alors d'assurer plus de validité et d'équité au processus d'orientation des élèves à l'entrée en 7<sup>e</sup>. En 2004, le système d'évaluation scolaire est entièrement revu. Le nouveau cadre général de l'évaluation prévoit des épreuves externes standardisées à différents degrés de l'école obligatoire. En 2005, celles-ci sont assurées à la fin de chaque cycle de deux ans: en 2<sup>e</sup>, en 4<sup>e</sup> et en 6<sup>e</sup>. En 2006, une autre épreuve externe est instaurée en fin de 8<sup>e</sup> année.

### *Informations communiquées aux directeurs d'école, aux inspecteurs, aux enseignants*

Dans tous les cantons qui en disposent, les résultats des élèves aux épreuves externes standardisées sont récoltés et analysés par les services concernés du département de l'instruction publique. Les analyses statistiques des résultats sont communiquées au différents niveaux de l'administration scolaire. A travers le genre d'analyses effectuées et les modalités de communication de ces analyses, on se rend bien compte de ce que les épreuves externes représentent pour l'administration scolaire en plus du rôle qu'elles jouent dans l'évaluation des élèves: une source d'indicateurs du rendement scolaire à l'échelle du canton, des établissements ou des classes. En d'autres termes, en référence à la figure 1, les épreuves externes contribuent au monitoring, c'est-à-dire à l'évaluation du système d'enseignement. A titre indicatif, voyons quels genres d'analyses statistiques des résultats aux épreuves externes sont communiqués aux inspecteurs ou aux directeurs d'établissements scolaires dans les cantons de Genève et de Vaud.

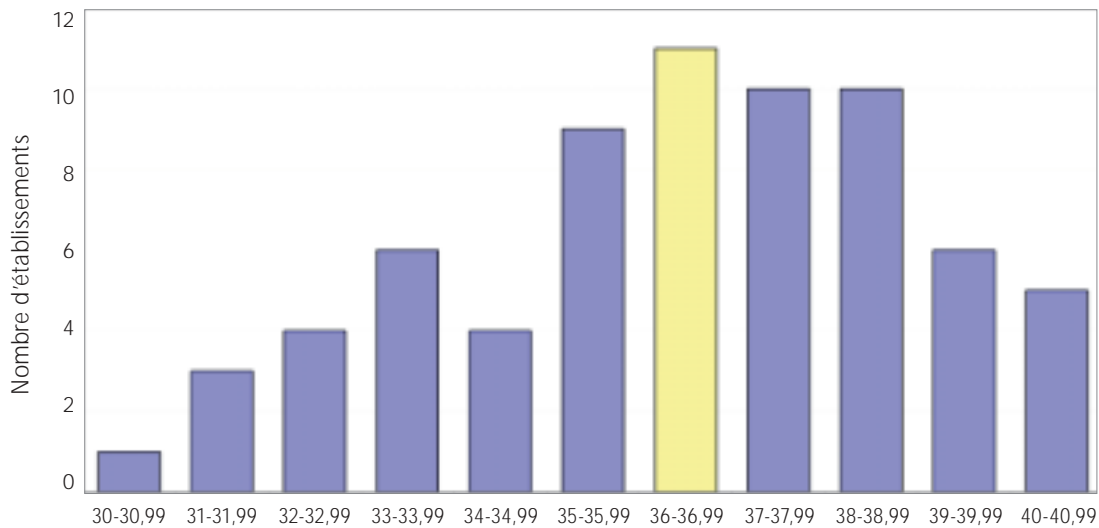
A l'école primaire du canton de Genève, deux indices statistiques du rendement des élèves sont communiqués aux responsables d'écoles ainsi qu'aux enseignants: le score moyen et le taux de réussite par épreuve, à l'échelle du canton et de l'école. Le rendement de l'école peut être ainsi comparé à celui du canton. Ces indices sont réunis dans un tableau (voir à titre d'exemple le tableau 5 concernant les résultats à l'épreuve externe de 6<sup>e</sup> passée en mai 2007) qui indique également pour chaque épreuve le score maximum ainsi que le seuil de réussite en fonction duquel les élèves sont évalués.

*Tableau 5: Canton de Genève. Tableau des résultats de l'école XX comparés aux résultats de l'ensemble du canton aux épreuves cantonales de 6<sup>e</sup> en mai 2007.*

	Français I	Français II	Mathématiques	Allemand
Score maximum	62	53	32	55
Seuil de réussite	42	33	21	36
Score moyen de l'école	XX	XX	XX	XX
Score moyen du canton	49.1	41.4	25.1	42.6
Taux de réussite de l'école	XX	XX	XX	XX
Taux de réussite du canton	84.3%	88.9%	82.3%	82.1%

Dans le canton de Vaud, la moyenne des résultats de l'ensemble des élèves du canton est communiquée aux parents. Les enseignants reçoivent en plus les distributions des résultats ainsi que les barèmes des notes. Les directeurs d'établissement reçoivent encore les moyennes des résultats de leur établissement ainsi qu'un histogramme de distribution des moyennes de tous les établissements du canton (voir, à titre d'exemple, la figure 2 concernant les résultats à l'épreuve de français passée en sixième en mars 2007). Chaque directeur d'établissement est ainsi en mesure de situer le niveau moyen des élèves de son établissement par rapport à celui de l'ensemble du canton et des autres établissements.

Figure 2: Canton de Vaud. Distribution des scores moyens des établissements (N=69) à l'épreuve cantonale de référence de français passée en sixième en mars 2007.



Résultats obtenus en moyenne – Moyenne cantonale: 36.3 sur 53 points

L'auto-évaluation des enseignants semble être la visée principale de la communication des statistiques du rendement des écoles. L'administration scolaire aide les enseignants à évaluer le résultat de leur travail en comparant les performances de leurs élèves à celles des élèves d'autres établissements et de l'ensemble du canton. Pour les inspecteurs et les directeurs d'établissement, les statistiques des résultats aux épreuves externes contribuent à leur faire connaître la situation des écoles dont ils ont la charge et s'avèrent utiles pour leur appréciation du travail des enseignants.

## Quelques perspectives pour les années à venir

Quittant le niveau cantonal, nous esquissons ici deux perspectives touchant d'une part au niveau régional – la Suisse romande – d'autre part au niveau national, via le projet HarmoS.

### *Les épreuves de référence romandes: un projet de la CIIP*

Nous avons vu que dans l'ensemble des cantons de Suisse romande des épreuves externes de référence sont conçues, répondant à des fonctions diverses. Dans tous les cas, cependant, elles donnent lieu à une passation auprès de tous les élèves du degré concerné.

Etant donné cette diversité et le fait que chacun, dans son canton, constitue un matériel propre, ne serait-il pas rationnel et utile de concevoir des épreuves romandes, centralisées et mises à la disposition des cantons, voire des enseignants? Cette idée n'est pas d'aujourd'hui, et à deux reprises la CIIP a mis au travail un groupe romand chargé de définir une conception commune de l'évaluation et de faire des propositions concrètes. Dans un premier temps, l'accent a été mis sur l'évaluation du travail de l'élève, en particulier l'évaluation formative. Mais, il s'est avéré que les processus d'évaluation scolaire sont liés et qu'il s'agit à la fois de les clarifier (buts et fonctions) et d'établir comment ils s'articulent entre eux.

Dans un second temps, les travaux entrepris dans ce cadre se sont déroulés au début des années 2000; en effet, la CIIP a mandaté un groupe de référence pour l'évaluation du travail des élèves (GRETTEL) en lui demandant de reformuler les lignes directrices de l'évaluation valables pour la Suisse

romande. Pourquoi parle-t-on ici de reformulation? Lors de la première opération, mise en place une dizaine d'années auparavant, le rapport qui en avait résulté (Weiss, 1996) présentait déjà une conception de l'évaluation scolaire et avait trouvé une suite dans les débats et dans des propositions concrètes de réformes lors d'un colloque romand sur l'évaluation (Weiss & Wirthner, 1998). Dix ans plus tard, il s'agissait alors de faire le point sur les réformes entreprises et de réaffirmer les lignes de force de l'évaluation du travail des élèves pouvant faire l'objet d'un consensus entre les cantons. Un nouveau rapport en a résulté et de nouvelles propositions (Weiss, 2004), parmi lesquelles figure la demande d'épreuves de référence romandes. Dans l'optique du GRETEL, ces épreuves devaient surtout avoir une fonction diagnostique. Sur la base de ce rapport, la CIIP a décidé de prendre deux mesures particulières, l'une précisément sur les épreuves de référence, l'autre sur les profils de compétences en fin de scolarité obligatoire.

A propos de la première mesure, voici quel en était l'objectif :

*« En accord avec les conclusions du rapport du Groupe de référence Evaluation du travail des élèves (GRETEL), l'intégration des instruments d'évaluation, permettant à l'enseignant de gérer, en cours de cycle, l'atteinte des balises par l'élève, doit être réalisée ; les moyens et supports d'enseignement se développeront en même temps ».*

La mesure elle-même est libellée de la façon suivante :

*« Création d'épreuves de référence romandes dans les principales disciplines du plan cadre en synergie avec les propositions du projet national HarmoS ».*

Cette formulation ouvre un champ de possibles plus vaste que les seules épreuves diagnostiques. Il reste donc encore à circonscrire ce champ et à délimiter quelles pourront être ces épreuves pour permettre de satisfaire les attentes et les demandes de l'ensemble des cantons. Dans cette perspective, la question de la confection d'une banque d'items est également posée. Ce projet est en cours de réalisation et devrait être confié à l'IRD.

### *Les standards de compétences dans le cadre du projet HarmoS*

Épreuves cantonales, épreuves romandes, épreuves internationales (de type PISA), les évaluations ne manquent pas ! A ces épreuves vont encore s'ajouter à moyen terme des épreuves à échelle nationale, si l'on en croit les propositions du projet HarmoS. Ce projet comporte plusieurs volets dont un relatif à la qualité de l'école suisse. C'est à l'intérieur de ce volet qu'est prévue l'évaluation du système sur la base de standards de formation et d'épreuves de mesure de l'atteinte de ces standards.

*Tableau 6 : HarmoS : monitoring du système scolaire suisse*

I. But / principes	Scolarité obligatoire : harmonisation – qualité – perméabilité – mobilité
II. Finalités	<b>Finalités</b> de l'école obligatoire : quelle formation de base pour tout élève en Suisse ?
III. Eléments	<b>Structures</b> : scolarisation – durée des degrés scolaires – aménagement de la journée scolaire
IV. Qualité	<b>Instruments du pilotage national</b> : standards (développement, détermination, évaluation, révision) – monitoring – plans d'études/moyens d'enseignement – portfolios
V. Autres	Adaptation structures cantonales et standards : 4 ans à partir de l'entrée en vigueur de l'accord

L'évaluation envisagée dans le cadre d'HarmoS doit donc fournir une aide au pilotage national des systèmes scolaires en vue d'une amélioration de leur qualité. A cette fin, plusieurs étapes sont définies : la formulation de standards de formation (de base) valables pour toute la Suisse, dans quatre disciplines (langue de scolarisation, mathématiques, langues étrangères et sciences naturelles), en fin de 2P, de 6<sup>e</sup> et de 9<sup>e</sup> année ; l'élaboration de plans d'études régionaux ; la confection d'épreuves pour vérifier l'atteinte des standards auprès d'échantillons d'élèves ; la régulation des systèmes en fonction des résultats obtenus (révision des plans d'études, des moyens d'enseignement, des portfolios de compétences, etc.).

Que faut-il entendre par standards de formation ? Ce sont des standards minimaux élaborés à partir d'un modèle de compétences préalablement décrit par des spécialistes des disciplines concernées. Ces standards :

- formulent des exigences à l'égard de l'enseignement et de l'apprentissage scolaires,
- désignent les objectifs du travail pédagogique en les exprimant en termes de résultats scolaires attendus.

Nous voulons souligner la caractéristique de la démarche adoptée en Suisse, de type *top down* ; en effet, le point de départ est l'élaboration, par des groupes de spécialistes, de modèles de compétences pour chacune des disciplines concernées, à partir d'une définition donnée d'emblée, celle de Klieme (2003), reprise par Weinert : « *Les capacités et aptitudes cognitives dont l'individu dispose ou qu'il peut acquérir pour résoudre des problèmes précis, ainsi que les dispositions motivationnelles, volitives et sociales qui s'y rattachent pour utiliser avec succès et responsabilité les résolutions de problèmes dans des situations variables.* » (Weinert, 2001, p.27). Une fois élaborés, ces modèles ont pris corps dans un ensemble de tâches testées auprès d'un échantillon d'élèves national ; cette passation d'épreuves a constitué la phase empirique du projet, assurée avec l'aide d'un groupe méthodologique. Sur la base des résultats obtenus selon le modèle statistique de Rasch, permettant une hiérarchisation des items selon leur difficulté, des niveaux de compétences ont été définis et ont servi à la formulation des standards de base. Ces standards devraient servir de référence pour la rédaction des plans d'études régionaux. Comme on peut le constater, des spécialistes (des disciplines et des statistiques) aux enseignants dans leur classe, une cascade d'étapes s'intercalent, avec, à notre sens, le risque d'un décalage entre les propositions des uns et les réalités des autres.

A la fin de l'année 2007, les groupes de travail des quatre disciplines concernées ont envoyé à la CDIP le modèle de compétences, les standards et un rapport court présentant leurs travaux. L'année 2008 sera consacrée à l'amélioration de ces documents en vue d'une large consultation dans les cantons, ainsi qu'à d'éventuelles études complémentaires dans certaines disciplines.

Au terme de cette présentation, la question de la multiplication des évaluations externes, à niveau cantonal, régional, national et international, se pose de façon aiguë. Il s'agirait pour le moins d'éviter des doublons et de gérer l'ensemble de telle manière que ces diverses évaluations apportent aux responsables scolaires, aux enseignants, aux élèves et au public, des informations utiles et complémentaires. Leur fonction, les modalités de passation et leur conception doivent donc être clairement explicitées.

## Questions et discussion

A la fin de chacun des ateliers, une discussion a eu lieu. Nous retraçons ici les principaux éléments de ces débats, sans qu'il soit possible d'être exhaustif – ce qui ne nous paraît d'ailleurs pas utile – mais en rassemblant les propos selon des thématiques dégagées à partir des notes que nous avons prises.

Les échanges entre les participants à notre atelier ont porté sur différents aspects de l'évaluation externe : ses visées, ses modalités, ses apports, ses limites et, évidemment, sa possible utilité pour l'évaluation des enseignants.

### *Epreuves externes et monitoring de l'enseignement*

La perspective d'évaluer la qualité de l'enseignement et des enseignants sur la base des résultats aux évaluations externes suscite, chez certains, scepticisme, méfiance et incompréhension. A leurs yeux, les résultats des élèves aux épreuves externes ne suffisent pas pour estimer la qualité de l'enseignement. A l'appui de cette affirmation, beaucoup d'intervenants soulignent le caractère fluctuant des résultats d'une école d'une année à l'autre. D'autres pensent que les résultats des épreuves externes arrivent souvent trop tard pour véritablement permettre une régulation de l'enseignement (les enseignants changent de classe, voire de degré d'enseignement...). Ainsi, les épreuves externes ne peuvent-elles être qu'un élément parmi d'autres à même d'assurer le monitoring de l'enseignement.

Cependant, chacun a conscience que la tendance actuelle au pilotage de l'enseignement par les résultats, d'origine anglo-saxonne, se développe dans nos contrées. Nous signalons que les fondements et les limites de ce modèle de gestion scolaire sont analysés dans un récent ouvrage collectif édité par Lessard et Meirieu (2005) sous le titre : *L'obligation de résultats en éducation. Evolution, perspectives et enjeux internationaux*.

Les discussions ont encore montré qu'il n'y a pas unanimité sur ce qui conviendrait le mieux à une régulation de l'enseignement : une évaluation diagnostique ou une évaluation bilan ?

### *Epreuves externes et évaluation des enseignants*

Il faut tout d'abord distinguer deux liens entre les résultats des épreuves externes et l'enseignement : les résultats peuvent constituer un indicateur du travail de l'enseignant ; ils peuvent apporter des informations utiles à l'enseignant et, de ce fait, être à son service.

Certains directeurs et inspecteurs de l'enseignement primaire sont d'avis que dans les deux cas, les épreuves externes peuvent contribuer à la régulation et à la qualité de l'enseignement. Par exemple, lorsqu'elles créent une certaine émulation chez les enseignants, qui, à travers elles, perçoivent plus clairement les objectifs à travailler, et cherchent à amener leurs élèves à les atteindre le mieux possible. Par ailleurs, ces épreuves fournissent également des informations utiles aux directeurs et inspecteurs, une base pour discuter avec certains enseignants de leur travail, en particulier lorsqu'il y a problème.

### *Epreuves externes standardisées ou banque d'items*

Certains voient dans les épreuves externes standardisées, du type de ce qui se prépare dans le cadre de HarmoS, un « joujou » pour technocrates, loin des réalités de l'enseignement. Dans ce cas, il leur semblerait préférable d'envisager la création d'une banque d'items, à la disposition des enseignants, permettant une réelle régulation de l'enseignement, une aide pour l'enseignant dans sa classe. La mise en place d'une banque d'items et son utilisation par les enseignants impliquent que l'institution scolaire fixe clairement les objectifs que les élèves doivent atteindre. Ainsi, les enseignants peuvent eux-mêmes vérifier si ceux-ci sont bien atteints par leurs élèves.

Il est à craindre encore que les épreuves externes soient une entrave à la liberté d'enseignement, en imposant des contenus et des modes de faire – conçus pour la mesure – que les enseignants se sentent contraints de travailler avec leurs élèves, au détriment peut-être d'autres savoirs ou savoir-faire.



### *Conception et concepteurs des épreuves externes*

Une autre question discutée a porté sur le profil des concepteurs des épreuves externes. Certains directeurs et inspecteurs de l'école primaire estiment que les épreuves externes sont appréciées par les enseignants parce qu'elles sont élaborées par des enseignants. A leur avis, ceux-ci se montreraient méfiants si l'élaboration des épreuves était confiée à des experts sans expérience d'enseignement. Or, il est prévu que les épreuves suisses, et romandes dans une certaine mesure, soient réalisées par des experts. Les épreuves internationales de type PISA le sont également. Pour d'autres, s'il est vrai que des épreuves conçues par des spécialistes non enseignants sont susceptibles de poser problème, des épreuves faites par des enseignants peuvent elles aussi se révéler problématiques. La standardisation des épreuves répond à des critères méthodologiques et techniques que seuls les experts en didactique des disciplines et les experts en évaluation sont en mesure de garantir.

Finalement, la solution se trouverait dans une collaboration entre experts et enseignants pour la réalisation des épreuves externes.

Un autre problème évoqué est celui des conditions de passation des épreuves. Dans la plupart des cas, ce sont les enseignants qui assurent cette passation dans leur classe, avec comme effet qu'elles varient en fonction des enseignants, ce qui provoque des biais dans les résultats obtenus.

En guise de conclusion, nous voulons souligner combien l'évaluation des enseignants par des épreuves externes reste une question difficile, sujette à controverse, voire ambiguë, à l'origine de craintes, telles que la mise en compétition des établissements ou la stigmatisation de certains enseignants. Les discussions ont montré, en tout cas, qu'il n'existait pas d'avis unanime sur cette question.

### Références bibliographiques

- De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris : PUF.
- Klieme, E. (Hrsg.) et al. (2003): *Zur Entwicklung nationaler Bildungsstandards: eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Lessard, Cl. & Meirieu, Ph. (éds). (2005). *L'obligation de résultats en éducation: évolution, perspectives et enjeux internationaux*. Bruxelles: De Boeck.
- Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen: eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (pp. 17-31). Weinheim/Basel: Beltz.
- Weiss, J. (1996). *Vers une conception cohérente de l'évaluation pour la scolarité obligatoire en Suisse romande et au Tessin: une évaluation pour apprendre et pour choisir: voies et moyens: rapport no 2*. Neuchâtel: IRDP.
- Weiss, J. (2003). *L'évaluation externe dans un concept général et cohérent d'évaluation: actes du 15e colloque international de l'ADMEE-Europe et congrès annuel de la SSRE, Université de Lausanne, 5-7 septembre 2002*. Lausanne: ISPPF.
- Weiss, J. (2004). *Lignes directrices pour un système d'évaluation du travail des élèves dans la scolarité obligatoire: rapport du GRETEL*. Neuchâtel: Conférence intercantonale de l'Instruction publique de la Suisse romande et du Tessin (CIIP).
- Weiss, J. & Wirthner, M. (éds). (1998). *Evaluation 97: pour une évaluation plus formative: actes du colloque, Yverdon-les-Bains, 20 novembre 1997*. Neuchâtel: IRDP.