QUELLE EVALUATION DES ACQUIS DES ELEVES POUR L'ECOLE GENEVOISE ? ETAT DES LIEUX ET REPRESENTATIONS DES ACTEURS

Anne Soussi, Christian Nidegger

Service de la recherche en éducation
12, Quai du rhône
CH-1205 Genève
Anne.soussi@etat.ge.ch
Christian.nidegger@etat.ge.ch

Mots-clés : *évaluation, représentations, pratiques*

Résumé. *Depuis un certain nombre d'années, l'évaluation externe a pris un essor considérable en Europe sur le modèle de l'acountability de l'Amérique du Nord et des pays anglo-saxons en général. En Suisse romande, le canton de Genève est celui qui a à la fois la plus longue tradition en matière d'évaluation externe et en compte le plus grand nombre. Ces évaluations externes participent, au cours de la scolarité obligatoire, à la certification des élèves, à côté de l'évaluation des enseignants.*

Après avoir rappelé brièvement quelles sont ou devraient être les finalités de l'évaluation d'après les enseignants, nous donnerons quelques éléments de leurs pratiques recueillies dans une enquête par questionnaire. Dans un deuxième temps, sur la base d'une autre étude, nous chercherons à mettre en évidence les représentations que les différents acteurs ont de cette évaluation. Quelles sont les relations entre les résultats aux évaluations externes à la fin de 6e, au début et à la fin du secondaire I et l'orientation des élèves? Quels sont les points forts et les points faibles du dispositif actuel? Nous nous intéresserons également aux liens entre les résultats aux évaluations externes et les résultats de l'année.

---

1. Contexte

Depuis un certain nombre d'années, l'évaluation externe a pris un essor considérable en Europe (cf. Eurydice, 2009) sur le modèle de l'acountability de l'Amérique du Nord et des pays anglo-saxons en général. Toutefois, l'évaluation peut avoir des visées et une place différentes selon les pays: être plutôt de type formative, diagnostique ou certificative, s'adresser aux élèves, aux enseignants, aux parents, aux autorités scolaires et porter sur les acquis des élèves, sur les établissement ou encore sur le système dans son ensemble. Parfois, elle joue plusieurs rôles à la fois comme le rappellent Demeuse et Monseur (2005).

En Suisse romande, le canton de Genève est celui qui a à la fois la plus longue tradition en matière d'évaluation externe et en compte le plus grand nombre. Toutefois, la majorité des cantons romands ont mis en place des épreuves de référence comme l'illustre le tableau suivant pour 2006-2007 (Ntamakiliro, 2010). Ces épreuves sont de formes variées et visent des buts différents selon les cantons et les degrés: évaluation diagnostique, formative, certificative; elles peuvent également concerner l'évaluation des élèves ou l'évaluation du système (pilotage ou monitorage).

Degrés	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

Berne									
Fribourg		●		●		●			
Genève		●		●		●	●	●	●
Jura						●			
Neuchâtel						●			
Valais				●		●		●	●
Vaud		●		●		●		●	

Figure 1: Epreuves de référence certificatives en Suisse romande en 2006-2007

A Genève, elles existent depuis relativement longtemps (introduites progressivement dans les années 90) mais ont évolué en se standardisant de plus en plus du point de vue de leur passation et de leur correction et en se multipliant. Si elles visent principalement l'évaluation des acquis des élèves, elles sont comme ailleurs utilisées de manière informelle à piloter le système comme le souligne Ntamakiliro (2010). Depuis quelques années, elles sont essentiellement de type certificatives, c'est-à-dire qu'elles participent à la certification des élèves en fin de cycles ou d'année, à côté de l'évaluation des enseignants. Ces évaluations s'effectuent annuellement, pour le primaire en 2e, 4e et 6e années en français et mathématiques ainsi qu'en allemand dès la 4e. Au secondaire I (cycle d'orientation), elles ont lieu en 7e, 8e et 9e années dans de nombreuses disciplines (français, mathématiques, allemand, anglais, biologie, etc.). A ces évaluations externes visant à évaluer les acquis des élèves, s'ajoutent d'autres formes d'évaluations externes s'intéressant davantage au pilotage du système comme l'évaluation PISA[1].

Dans cette présentation, après avoir rappelé brièvement quelles sont ou devraient être les finalités de l'évaluation d'après les enseignants, on donnera quelques éléments de leurs pratiques recueillies dans une enquête par questionnaire en 2005. Dans un deuxième temps, on cherchera à mettre en évidence les représentations que les différents acteurs ont de cette évaluation. Quelles sont les relations entre les résultats aux évaluations externes à la fin de 6e, au début et à la fin du secondaire I et l'orientation des élèves? Quels sont les points forts et les points faibles du dispositif actuel? On s'intéressera également aux liens entre les résultats aux évaluations externes et les résultats de l'année. Quelle est la place des différentes évaluations (internes ou externes au niveaux cantonal, régional, national ou encore international)?

2. Méthodologie

Dans cette présentation nous nous appuierons essentiellement sur les données recueillies provenant de deux études. La première est une enquête par questionnaire portant sur les pratiques déclarées d'évaluation à l'école obligatoire (et au secondaire II) réalisée en 2005 auprès de tous les enseignants des différents niveaux (Soussi et al., 2006). La seconde source provient d'une étude effectuée en 2008 (Soussi, Guilley, Guignard et Nidegger, 2009) dans le cadre d'une réflexion visant à réorganiser la scolarité obligatoire en regroupant notamment les deux niveaux d'enseignement, primaire et secondaire I. Dans cette étude, un état de situation des dispositifs existants de l'évaluation cantonale commune a été réalisé au moyen d'entretiens et d'un bref questionnaire auprès des différents acteurs concernés (autorités scolaires, concepteurs d'épreuves

[1] Depuis la première enquête PISA en 2000, tous les cantons romands participent à cette enquête avec un échantillon complémentaire d'élèves de 9e année permettant de disposer des résultats de cette enquête au niveau de chaque canton.

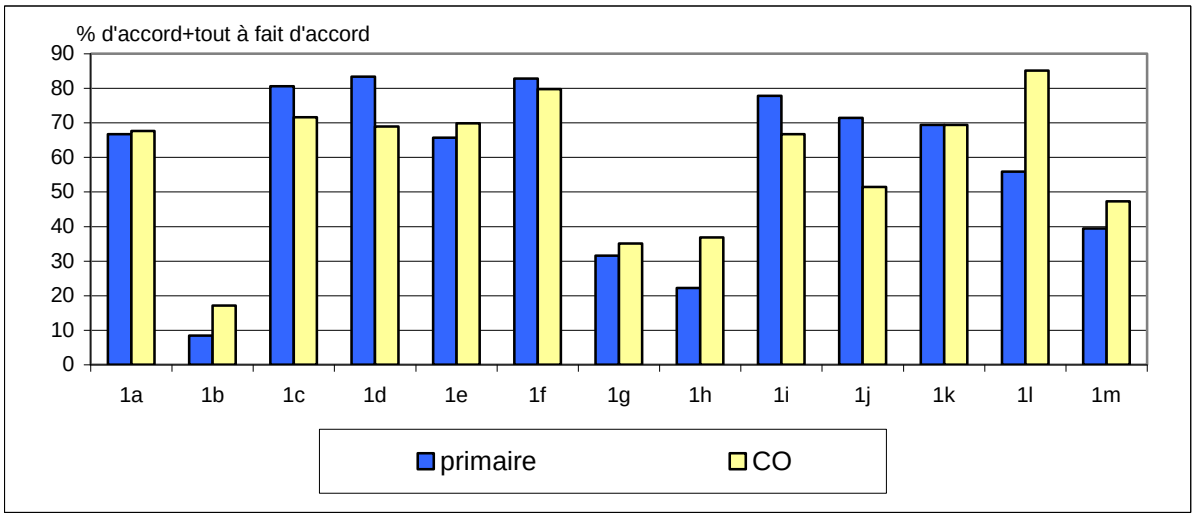
et un échantillon d'enseignants du primaire et du secondaire I). Par ailleurs, des analyses secondaires des résultats aux épreuves cantonales ou communes ont également été faites pour mettre en évidence les parcours d'élèves sur les trois années du secondaire I (cycle d'orientation) ainsi que des analyses qui mettent en évidence les liens entre les notes données par les enseignants et les résultats aux épreuves cantonales (école primaire).

3. Résultats

Les enseignants interrogés lors de la première enquête (2005) estimaient que l'évaluation vise surtout à vérifier les acquis et l'atteinte des objectifs, à situer les élèves dans un processus d'apprentissage et à leur donner des repères et de façon un peu moins marquée à orienter, réguler et comprendre les démarches des élèves. On retrouve là les deux fonctions les plus fréquentes de l'évaluation: sommative/certificative et formative. Par ailleurs, l'évaluation doit être intégrée au processus d'enseignement. La conception varie entre les deux niveaux d'enseignement (primaire et secondaire I). Les enseignants du primaire estiment de manière plus élevée que leurs collègues du secondaire I que l'évaluation doit porter surtout sur ce qui a été travaillé en classe et qu'il est important de chercher à savoir pourquoi un élève ne réussit pas. Pour ceux du secondaire I, mettre des notes est une finalité importante. On peut supposer que ces différences de centration sont fonction des enjeux prioritaires aux différentes étapes du cursus scolaire.

D'autres écarts apparaissent également au niveau des pratiques déclarées d'évaluation et de leur fréquence. Les instruments utilisés pour évaluer les élèves sont avant tout écrits mais au primaire, d'autres instruments (définis alors dans les instructions officielles) viennent les compléter tels que les porte-folios ou les observations d'élèves en situation. Quant à la fréquence des évaluations, elle est variable: les épreuves portant sur un champ plus large sont pratiquées par 2/3 des enseignants des deux niveaux au moins deux fois par période, par contre les évaluations portant sur un champ plus restreint sont plus fréquentes au primaire qu'au secondaire I d'après les enseignants interrogés. Il en va de même pour l'évaluation formative. L'évaluation sommative est par contre un peu plus fréquente au secondaire I. On peut faire l'hypothèse que ces écarts sont dus à l'organisation différente des deux niveaux d'enseignement: d'un côté un enseignant généraliste et de l'autre un enseignant spécialiste d'une discipline confronté à un découpage horaire strict.

En plus de l'évaluation réalisée par les enseignants, les élèves sont soumis à une évaluation certificative externe. On observe chez les enseignants un certain nombre de similitudes entre les objectifs de ces deux types d'évaluation (interne et externe). Les enseignants du primaire comme ceux du cycle d'orientation attribuent aux épreuves les deux principaux objectifs annoncés par l'institution, à savoir vérifier si les élèves ont acquis les objectifs attendus et réguler l'enseignement. A cela, ils ajoutent parfois une dimension de pilotage et de monitorage. Au cycle d'orientation, la dimension d'orientation est mise en évidence ce qui n'est pas le cas au primaire même en 6e, moment du passage au cycle d'orientation et de la répartition des élèves dans les filières en fonction de leurs résultats scolaires. Par ailleurs, leur opinion sur l'évaluation externe diffère peu comme l'illustre le graphique suivant, sauf pour trois items: *L'évaluation externe entraine un certain bachotage* (1l) et *L'évaluation externe oblige (ou a pour effet) de limiter le champ de l'enseignement au contenu des épreuves (ou de l'évaluation)* (1m) ou encore *L'évaluation externe ne tient pas compte de ce qui est vraiment enseigné en classe* (1h). Pour les deux premières affirmations, les enseignants du secondaire I (CO) ayant répondu font preuve d'un accord sensiblement plus élevé. Pour la troisième, si les enseignants ont plutôt tendance à ne pas être d'accord avec cette affirmation concernant le rapport entre ce qui est enseigné en classe et ce qui est évalué dans l'évaluation externe, ceux du CO sont près de 40% à être plutôt d'accord voire tout à fait d'accord (alors que ce n'est le cas que d'un enseignant sur 5 au primaire).



a) L'évaluation externe peut être génératrice d'idées nouvelles.
b) L'évaluation externe n'est pas pertinente car seul l'enseignant est à même d'évaluer les apprentissages de ses élèves parce qu'il les connaît bien.
c) L'évaluation externe contribue à harmoniser les pratiques d'évaluation entre enseignants.
d) L'évaluation externe permet une certaine objectivité parce qu'elle tient compte de ce qui est attendu pour un degré ou un cycle donné.
e) L'évaluation externe permet d'évaluer et de réguler le système.
f) Il est important d'avoir, à côté de l'évaluation régulière réalisée par l'enseignant, une évaluation commune à tous les élèves d'un degré ou pour un cycle car elle donne des repères précis et objectifs.
g) L'évaluation externe sert à évaluer les enseignants.
h) L'évaluation externe ne tient pas compte de ce qui est vraiment enseigné en classe.
i) L'évaluation externe est utile pour s'assurer que les objectifs fondamentaux sont atteints.
j) L'évaluation externe permet de savoir précisément ce qui est attendu.
k) L'évaluation externe a pour effet de stresser les élèves.
l) L'évaluation externe entraine un certain bachotage.
m) L'évaluation externe oblige (ou a pour effet) de limiter le champ de l'enseignement au contenu des épreuves (ou de l'évaluation).

Figure 2: Opinion sur l'évaluation externe (au primaire et au secondaire I)

Par ailleurs, les informations recueillies auprès des différents acteurs (autorités scolaires, concepteurs des épreuves, enseignants), lors d'une deuxième enquête réalisée en 2008, montrent qu'il existe un consensus, en tout cas formel, sur les objectifs des épreuves externes genevoises aussi bien pour l'enseignement primaire que pour le cycle d'orientation. Elles visent avant tout la vérification de l'atteinte des objectifs du programme et elles entrent en ligne de compte pour la certification des élèves. Cependant dans la pratique, les épreuves sont parfois utilisées à d'autres fins (pilotage du système, par ex.) pour lesquels elles n'ont pas été conçues. Par exemple, même si les résultats ne sont pas publics, les directions des ordres d'enseignement disposent de résultats permettant la comparaison des moyennes de classes ou d'établissements. De plus, derrière cette cohérence formelle, il faut être toutefois attentif aux spécificités liées aux niveaux d'enseignement, aux parcours scolaires et à l'âge des élèves: par exemple, l'utilisation d'une épreuve commune de type papier-crayon entrant dans la certification pour des élèves de 2P. A cet âge, on peut se demander si l'aspect diagnostique ne devrait pas être privilégié. Pour les épreuves de 6P, si elles

visent officiellement la vérification du programme, elles sont également utilisées pour l'orientation des élèves au cycle d'orientation.

4. Apports et limites des évaluations externes

Nous avons également tenté d'identifier, par rapport à cette évaluation externe, quels étaient les points forts et les points faibles du dispositif actuel repérés par les groupes d'acteurs en utilisant trois critères : l'utilité, la qualité technique et l'efficacité/efficience (cf. Stufflebeam, 1974, 2003; Scriven, 1974; Stake, 1969).

4.1 Utilité des épreuves cantonales ou communes

Pour la plupart des personnes interrogées dans les deux niveaux d'enseignement, un des principaux points forts de ces évaluations est d'être commune à tous les élèves d'une cohorte. Cela permet d'avoir les mêmes exigences pour une discipline donnée. Elles permettent également de réguler l'enseignement et d'unifier les pratiques. Elles peuvent aussi contribuer à une certaine crédibilité par rapport aux niveaux d'enseignement suivants (cycle d'orientation ou secondaire II). Elles sont également importantes pour les élèves et les parents. Les enseignants du cycle d'orientation relèvent également que les épreuves permettent d'habituer les élèves à être évalués sur un champ plus large et à fonctionner de manière autonome.

Pour les points faibles mis en évidence par les acteurs il ressort que ces épreuves sont utilisées pour répondre à trop de besoins: il y aurait notamment une confusion entre les objectifs visant à vérifier l'atteinte du programme par les élèves et le pilotage du système. Ainsi, les acteurs semblent avoir une certaine conscience de la difficulté ou du risque que les épreuves externes visent plusieurs objectifs à la fois. Pour d'autres, il manque une évaluation système (notamment au primaire). Par ailleurs, l'analyse et l'exploitation des résultats jugées insuffisantes sont soulignées par la majorité des acteurs. Certains effets pervers des épreuves sont relevés notamment au cycle d'orientation: la tendance au bachotage et la réduction du programme et de l'enseignement au contenu des épreuves. Certains enseignants du primaire déplorent la trop grande importance de ces épreuves aux yeux des élèves et de leurs parents ainsi que le stress engendré, en particulier chez de jeunes élèves. Enfin, au cycle d'orientation, on estime que le type d'évaluation proposé est trop différent de ce qui se fait habituellement en classe (notamment pour le regroupement B, filière à exigences faibles).

4.2 Qualité technique des épreuves cantonales ou communes

De manière générale, les différents acteurs interrogés (autorités, directions, concepteurs et enseignants) relèvent une bonne qualité technique des épreuves. On notera cependant quelques spécificités qui se rapportent plus souvent à un niveau d'enseignement en particulier.

Ainsi, au primaire, les différents acteurs soulignent l'existence de prétests, de tables de spécification qui permettent notamment d'identifier les objectifs évalués. Certains enseignants relèvent également le caractère objectif de ces épreuves, comparativement à celles élaborées par les enseignants pour leurs élèves. Les consignes de passation et de correction sont également jugées comme bien expliquées.

Enfin, l'expérience (et la stabilité) des concepteurs d'épreuves est également considérée comme un point positif. Il faut souligner que le dispositif comprend à la fois des formateurs/didacticiens des disciplines concernées et des formateurs en évaluation, l'ensemble étant encadré par le responsable des épreuves cantonales, spécialiste en évaluation.

Au cycle d'orientation, d'autres éléments sont mis en évidence et sont dépendants des disciplines considérées: une certaine stabilité de contenu et du niveau de difficulté, de bonnes conditions de passation et de correction, une certaine clarté de l'épreuve et des consignes de correction

(notamment en anglais), une bonne couverture du programme (par exemple en anglais et en physique) on encore un bon équilibre à l'intérieur des épreuves pour la plupart des disciplines.

Certains points faibles, concernant la qualité technique des épreuves, observés dans le dispositif actuel sont communs aux deux niveaux d'enseignement: des différences au niveau des approches selon les disciplines (et la nécessité d'harmoniser ces approches), une fiabilité insuffisante pour évaluer des connaissances ou compétences détaillées (par exemple en 2P ou dans une épreuve de français au cycle d'orientation). De manière générale, l'analyse des résultats voire le recours insuffisant (ou leur utilisation) aux prétests est également considéré comme un point faible aux deux niveaux. Par ailleurs, les prétests ne permettent pas pour l'instant d'évaluer la qualité technique des épreuves (par. ex. validité, fiabilité).

Les autorités scolaires relèvent encore un certain manque de standardisation des épreuves et souhaitent des équipes pluridisciplinaires avec des compétences aussi bien des curricula que de la docimologie, ainsi qu'une plus grande cohérence entre les deux niveaux d'enseignement.

4.3 Efficacité/efficience des épreuves cantonales ou communes

Ce dernier critère s'est avéré plus délicat à estimer par les différents acteurs interrogés. Parmi les points positifs, il a été relevé d'une part, une bonne prédictivité des épreuves et d'autre part, au cycle d'orientation en particulier, l'existence d'un outil informatique très précieux, EVACOM qui permet la saisie des résultats aux épreuves cantonales ou communes (informations récoltées pouvant être les plus complètes possibles selon les disciplines) et leur consultation.

Par contre, la place des épreuves cantonales ou communes dans les carnets et leur prise en compte de manière plus globale est considérée comme un problème par l'ensemble des acteurs. Au cycle d'orientation, certains enseignants relèvent la surcharge de travail due à la correction et à la saisie des résultats.

5. Liens entre évaluation interne des enseignants et évaluation externe

Les enseignants ont également été interrogés sur les liens entre l'évaluation externe et leur évaluation de l'année. A l'école primaire, les résultats aux épreuves cantonales correspondent, pour la plupart des domaines, à ceux de l'année, et ce pour la majorité des élèves voire seulement pour une partie. On observe une certaine variation entre disciplines. Ainsi, pour le français II (structuration), on observe plutôt que les épreuves cantonales confirment les résultats de l'année, il en va de même pour les maths en 2P. Pour l'allemand, notamment en 6P, l'épreuve qui était considérée comme facile, donne lieu à de meilleurs résultats que les évaluations durant l'année. Pour le français enfin, cela parait variable d'une école à l'autre, dépendant du quartier et de la composition socioéconomique ou linguistique du public d'élèves.

Au cycle d'orientation, s'il est difficile de mettre en évidence des tendances, on relèvera que le plus souvent les résultats aux épreuves communes correspondent à ceux de l'année pour une bonne partie des élèves de A et pour une partie des élèves de B. Quand on pose la question de savoir si les résultats aux épreuves communes confirment, sont meilleurs ou sont moins bons que les résultats aux évaluations de l'année, les réponses sont très contrastées selon les disciplines: en français et mathématiques, la tendance est qu'ils confirment ceux de l'année pour les élèves de A (filière à exigences élevées) et qu'ils sont meilleurs en mathématiques et en français pour les élèves de B (filière à exigences plus limitées[2]). C'est l'inverse en langues 2 et en physique pour les élèves de B. On voit bien là la difficulté d'avoir pratiquement les mêmes épreuves pour des

[2] Le système genevois actuel au secondaire I se compose de deux types d'établissements: 17 (sur 20) possèdent des regroupements différenciés A et B: le regroupement A ayant des exigences élevées et le regroupement B des exigences moins élevées et des effectifs réduits. Trois établissements comportent des classes hétérogènes en 7e ainsi que des niveaux (mathématiques et allemand) et des options.

regroupements avec un niveau d'exigence différent comme c'est surtout le cas en français et en anglais.

6. Liens entre résultats aux épreuves cantonales, orientation des élèves et notes des enseignants

Dans le cadre de cette étude, nous nous sommes également intéressés aux liens existant entre les résultats aux épreuves cantonales et l'orientation des élèves en observant le parcours des élèves qui ont passé les épreuves cantonales à la fin du primaire (6P) au cours des trois années suivantes. Même si les évaluations externes ne sont pas censées déterminer à elles seules l'orientation des élèves, on observe que 80% des élèves ayant atteint le seuil de réussite dans les trois domaines déterminants pour l'orientation se retrouvent dans la filière la plus exigeante (regroupement A). Si l'on observe la situation deux ans plus tard (9e année), on trouve, assez logiquement, un taux un peu plus bas de 69%. A l'opposé, on constate que les redoublement sont plus fréquents chez les élèves qui n'avaient réussi qu'un des trois domaines pris en compte dans l'orientation (français I, compréhension/production, français II, structuration et mathématiques). Globalement, les évaluations externes semblent avoir une assez bonne prédictivité pour l'orientation des élèves au cycle d'orientation même au bout de trois ans.

Il a également été possible de comparer les liens entre résultats aux épreuves externes et les évaluations annuelles réalisées par l'enseignant pour les trois degrés primaires (2, 4 et 6P). Si l'on observe les corrélations entre les deux types de mesures, on constate que les corrélations sont élevées (plus de 70%) et significatives statistiquement. Ces constats semblent être en contradiction avec les propos de certains enseignants interrogés dans une étude sur l'évaluation des acquis des élèves à Genève (Soussi, Guilley, Guignard, Nidegger, 2009) qui déclaraient que les évaluations externes ne correspondaient pas forcément à leurs évaluations de l'année dans un sens ou dans un autre selon la composition socioéconomique de la classe. En effet, si les enseignants interrogés dans cette étude pouvaient avoir cette impression, les corrélations calculées montrent que cette impression n'est pas confirmée par les analyses effectuées.

7. Place des différentes évaluations internes et externes et discussion

A côté de l'évaluation interne réalisée par les enseignants, coexistent et sont en train de se développer différentes formes d'évaluation externe: les épreuves cantonales ou communes destinées à tous les élèves d'une volée pour vérifier le degré d'atteinte des objectifs du programme, les épreuves romandes de référence (en cours d'élaboration) qui se baseront sur le plan d'études romand, les futurs tests nationaux basés sur les standards nationaux visant le monitorage du système et enfin, l'enquête internationale PISA visant également le pilotage et monitorage du système. La situation est actuellement en pleine mouvance. Certaines de ces évaluations externes visent le niveau de connaissance ou de compétences des élèves (évaluation interne, épreuves cantonales ou communes), d'autres plutôt l'évaluation du système (PISA, HarmoS) mais pour l'instant, il est difficile de savoir quelle fonction remplira l'évaluation romande (niveau de connaissance/compétence des élèves, validation ou vérification de l'atteinte du plan d'études romand), si elle sera diagnostique ou certificative. L'autre question à se poser est également celle du devenir des épreuves cantonales ou communes quand le plan d'études romand sera mis en place et remplacera le plan d'études cantonal.

Ces différentes évaluations internes et externes se complètent et ont chacune leur utilité. Toutefois, l'évaluation externe qu'elle vise l'évaluation des élèves et la régulation de l'enseignement comme le pilotage du système prend actuellement un essor considérable à Genève tout comme dans l'ensemble de la Suisse romande et des pays européens. Cette tendance est renforcée par le développement de l'autonomie des établissements et le besoin de rendre compte aux autorités scolaires. Il faut néanmoins garder en mémoire que le temps scolaire n'est pas extensible et qu'un équilibre doit être maintenu entre enseignement et évaluation. De plus, si l'évaluation externe

visant l'évaluation des élèves permet d'avoir des exigences communes mais plus difficilement de réguler directement l'enseignement et de faire progresser les élèves. C'est pourquoi, il parait nécessaire de conserver une évaluation plus formative pour identifier progrès et lacunes des élèves.

Par ailleurs, en plus de la définition de la place respective des différentes évaluations, une question épineuse reste posée: celle de mise en oeuvre d'un dispositif d'évaluation externe cantonal cohérent pour l'ensemble de la scolarité obligatoire. Les résultats de la 2ème étude ont montré l'existence de différences de points de vue aussi bien entre les différents acteurs qu'entre niveaux d'enseignement.

De plus, la réflexion sur l'évaluation externe devrait également tenir compte de l'importance donnée à cette évaluation par rapport aux évaluations des élèves réalisées par les enseignants. Plus globalement, la place de l'évaluation par rapport au temps consacré à l'enseignement et l'apprentissage doit également faire partie de la réflexion plus générale sur les objectifs de l'école.

8. Bibliographie

Eurydice (2009). *Les évaluations standardisées des élèves en Europ: objectifs, organisation et utilisation des résultats.* Bruxelles: Eurydice.

Monseur, Ch., Demeuse, M. (2005). Les évaluations externes permettent-elles une évaluation efficace? In M. Demeuse, A. Baye, M.H. Straeten, J. Nicaise & A. Matoul (Eds). Vers une école juste et efficace. Bruxelles: de Boeck.

Ntamakiliro, L., Tessaro, W. (2010). Plus d'épreuves externes: vers une standardisation de l'évaluation scolaire. In P. Gilliéron et L. Ntamakiliro (eds). Réformer l'évaluation scolaire: une mission impossible?

Scriven, M. (1974). Evaluation perspectives and procedure. In W.J. Popham (Ed). Evaluation in education. Berkeley: McCutchan.

Soussi, A. et al. (2006). EVALEPCOPO. Pratiques d'évaluation: ce qu'en disent les enseignants (à l'école obligatoire et dans l'enseignement obligatoire général. Genève: Service de la recherche en éducation.

Soussi, A., Guilley, E., Guignard, N, Nidegger, Ch. (2009). Evaluation des acquis des élèves à l'école obligatoire. Genève: Service de la recherche en éducation.

Stake, R.E. (1969). Evaluation, design, instrumentation, data collection and analysis of data. In J.L. David (Ed). Educational evaluation. Colombus, OH: State Superintendent of Public Instruction.

Stufflebeam, D.L. (1974). Metaevaluation. Occasional Paper Series, no 3. Kalamazoo, MI: The Evaluation Center, Western Michigan University.

Stufflebeam, D.L. (2003). Professional Standards and Principles for Evaluations. In International Handbook of Educational Evaluation, 279-302.