Étude de l'énoncé d'une épreuve écrite, dans le cadre de la didactique générale

Analyse contextualisée d'énoncés d'évaluations internes dans les disciplines scientifiques enseignées au collège

Résumé : L'objectif de cette communication est de montrer comment dans le cadre de la continuité didactique présente dans une classe entre les temps d'enseignement-apprentissage et ceux d'évaluation, des concepts de la didactique générale peuvent être utilisés par un professeur pour analyser la qualité de l'énoncé de l'épreuve qu'il a conçu pour ses élèves. Le cadre de cette recherche se limite aux disciplines scientifiques enseignées au collège.

Au cours d'une situation d'évaluation formelle à support écrit, le processus évaluatif se développe à travers des interactions (évaluatives) entre le professeur et les élèves, et chaque élève, en fonction des particularités de la classe. L'énoncé de l'épreuve est le principal objet à partir duquel ces interactions se développent. Il est en effet conçu par le professeur et reçu par chaque élève, en fonction de la réalité de la classe : ce qui est enseigné (le curriculum réel pour le professeur, réalisé pour chaque élève), les règles locales mises en place (le contrat didactique tel qu'il est compris par le professeur, et par chaque élève), les travaux menés en classe (les tâches prescrites par le professeur et les activités développées par chaque élève). Ainsi ces trois déterminants (curriculum, contrat didactique, tâches), qui sont issus de la didactique générale, d'une part sont des outils dont dispose tout professeur pour lui permettre d'explicitier ses intentions évaluatives, et d'autre part leur perception par les élèves influence la performance réalisée par chacun d'entre eux, ayant éventuellement pour conséquences, des inférences évaluatives partiellement erronées.

Mots-clé : situation d'évaluation formelle ; énoncé d'épreuve ; curriculum ; contrat didactique ; tâches prescrites

Introduction

L'énoncé d'une épreuve d'évaluation interne à support écrit est au centre du processus évaluatif. C'est en effet dans l'énoncé que se concrétisent les intentions évaluatives du professeur, et c'est lui qui va déclencher le comportement que chaque élève va avoir au cours de l'épreuve.

Dans le cas d'une évaluation interne, le professeur conçoit un énoncé non seulement en conformité avec le programme officiel (curriculum prescrit), mais surtout en fonction de la réalité conjoncturelle de sa classe (Mottier-Lopez, 2006). Les éléments qui constituent cette réalité sont notamment d'ordre sociologique (Merle, 2018), psychologique (Mialaret, 2011) ou encore didactique. En se limitant dans cette communication au cadre didactique, le curriculum (De Landsheere, 1979), le contrat didactique (Sarrazy, 1995) et les tâches prescrites sont des éléments constitutifs de la réalité de la classe, à une période donnée. De plus, ils sont bivalents, c'est-à-dire qu'ils ont un sens pour le professeur et pour chaque élève : le curriculum est réel pour le professeur et réalisé pour chaque élève ; le contrat didactique définit ce que le professeur peut attendre de chaque élève et réciproquement ; les tâches prescrites déclenchent les activités des élèves.

Dans cette communication, je me propose de présenter comment ces outils issus de la didactique générale (curriculum, contrat didactique, tâches prescrites) permettent à un professeur d'analyser contextuellement la qualité de l'énoncé qu'il a conçu pour évaluer ses élèves. Le champ est limité aux disciplines scientifiques enseignées au collège¹. Dans la perspective interactionniste de l'évaluation, la qualité de l'énoncé qui lui-même reflète les intentions évaluatives du professeur, est analysée à travers l'appréhension qu'en a chaque élève.

Problématique

Dès lors que l'évaluation est considérée comme un processus (De Ketele, 1993), il est possible théoriquement d'en fixer le commencement et la fin, et d'en décrire le déroulement. Nous désignons par situation d'évaluation (Raulin, 2017), l'intégralité du temps durant lequel le processus évaluatif se développe. Elle peut se décomposer en cinq phases : 1) la décision prise par le professeur d'organiser une épreuve, l'information qu'il en donne à ses élèves ; 2) l'élaboration de l'énoncé ; 3) l'épreuve ; 4) la correction des travaux des élèves, l'établissement du jugement évaluatif ; 5) la restitution aux élèves, la correction en classe. Une situation d'évaluation est qualifiée de formelle quand les élèves savent dès le commencement qu'ils seront évalués.

Habituellement, la phase 3 est considérée comme la plus importante parce que les élèves y réalisent leurs performances à partir desquelles le professeur va formuler son jugement évaluatif (note, niveaux de compétences, commentaires, annotations). Dans cette phase, le déclencheur est l'énoncé qui a été conçu par le professeur. C'est dans la phase 3 que les interactions qui se développent tout au long du processus évaluatif sont les plus manifestes.

L'énoncé, un objet élaboré sous contraintes.

Le professeur est en principe, contraint par le programme officiel, quand il rédige son énoncé, mais en fait il s'agit d'une référence plus théorique que réelle : en fait, le programme cadre plutôt les phases d'apprentissage en amont de l'évaluation. Plus que le programme officiel, c'est le programme tel qu'il est traité en classe qui sert de référence : pour le professeur, c'est le curriculum réel, pour un élève, c'est le curriculum réalisé (Perrenoud, 1993). Quand le professeur définit le cadre de la révision pour une situation d'évaluation formelle particulière, c'est au curriculum réel qu'il se réfère.

La deuxième contrainte tient au contrat qui régule les interactions évaluatives (Reuter, 2008) : le contrat didactique. D'après Brousseau (1997), il définit en effet ce que le professeur peut attendre de ses élèves, et inversement, ce que les élèves peuvent attendre de leur professeur.

La troisième contrainte relève de la construction même de l'énoncé : la répartition entre tâches de restitution, d'application ou nécessitant éventuellement une réflexion supplémentaire.

Problématique et question de recherche

La préoccupation initiale tient à la vacuité des appréciations portées sur la difficulté d'une épreuve d'évaluation interne quand elles sont énoncées sans tenir compte du contexte spécifique que représente le microcosme d'une classe à un moment donné (Mottier Lopez, 2006). La problématique porte sur la façon dont les interactions évaluatives se développent au cours d'une épreuve, entre les deux acteurs génériques d'une situation évaluative interne, le professeur et chaque élève : pour ce faire, symétriquement aux intentions évaluatives du professeur étudiées par Jorro (2006) et Issaieva (2013), il a fallu prouver l'existence de la conscience évaluative individuelle de chaque élève: « La conscience évaluative correspond à l'interprétation qu'un élève peut avoir des intentions évaluatives de son professeur, sa compréhension et son acceptation de celles-ci, avant d'entrer dans la phase active de réalisation.» (Raulin, 2017).

Après avoir montré que la conscience évaluative d'un élève dépend de paramètres conjoncturels comme les contenus d'enseignement et de l'enseignement ainsi que de la réalité du contrat didactique au moment considéré, la question de la compréhension des intentions évaluatives du

professeur par chaque élève est explorée et conséquemment celle du différentiel entre les intentions évaluatives du professeur et leur appréhension par chaque élève.

Cette réalité offre la possibilité aux professeurs d'expliciter leurs intentions évaluatives, s'ils le souhaitent.

Méthodologie

Corpus de données

L'objectif étant de répondre à une question d'existence, une méthode par étude de cas est la plus adaptée (Grawitz, 2001).

Dix études de cas ont été menées ; 105 élèves et 9 professeurs (2 de mathématiques, 2 de physique-chimie, 2 de SVT et 3 de technologie) provenant de 3 collèges situés dans deux académies différentes y ont contribué.

Le corpus de données est principalement constitué des réponses données à des questionnaires¹, par les professeurs et leurs élèves, à la suite d'une épreuve écrite. Ces questionnaires sont conçus de la manière suivante :

- ils sont composés de deux parties : la première porte sur une approche globale de l'énoncé (sa longueur, le nombre de questions, la durée de l'épreuve...), la deuxième suit l'organisation de l'énoncé en interrogeant sur chaque exercice, chaque question ; ils comportent en moyenne de l'ordre de 80 items ;
- pour analyser la perception des élèves et les intentions du professeur pour chaque question, des propositions portent sur « la présence de la réponse dans le cahier », « le fait d'utiliser ce qui a été fait en classe », « c'est une question où ce qu'on a fait en classe ne suffit pas » ;
- ils sont composés d'affirmations sur lesquelles on demande aux élèves et aux professeurs de donner leur niveau d'adhésion, selon une échelle de Likert à 4 niveaux² ;
- les propositions sont rédigées de façon que le professeur et les élèves aient pour répondre une référence commune (Grawitz, 2001), constituée de leur vécu commun de la classe.

Exploitation des données

L'exploitation des données brutes recueillies dans les questionnaires passe par la transformation de celles-ci en données chiffrées : pour ce faire, à chaque niveau de l'échelle de Likert est attribuée une cotation de 1 à 4³. Ces cotations sont ensuite additionnées pour obtenir un score (partiel ou global) qui représente ce qu'a répondu celui ou celle qui a rempli le questionnaire. Il est à noter que le choix des cotations ne permet pas que des compensations s'opèrent entre les items d'une part, et d'autre part que les cotations globale ou partielles n'ont aucune valeur intrinsèque (Grawitz, 2001).

En revanche, il est possible de les comparer : deux scores égaux peuvent correspondre à des réponses identiques, et deux scores différents sont l'assurance de réponses différentes.

1 En plus des questionnaires, ont été exploités marginalement des entretiens semi-directifs auprès des professeurs et de quelques-uns de leurs élèves, une observation d'épreuve, un tableau de bord de l'évaluation en ligne, tenu par des professeurs

2 Les quatre réponses proposées sont tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord. Il faut y ajouter systématiquement la possibilité de ne pas répondre : en cochant la case NSP (ne se prononce pas).

3 La cotation 0 correspond à une non-réponse.

Pour chaque étude de cas, les données recueillies ont donné lieu à ce travail de cotation, puis de comparaison des scores (partiels ou global) entre les réponses des élèves et celles du professeur.

Le tableau ci-dessous présente différents exemples extraits des données collectées à partir de questionnaires utilisés, à la suite d'un contrôle de SVT, en 4ème (23 élèves).

- les cotations 1 et 2 correspondent à un désaccord avec la proposition : 1, pas du tout d'accord ; 2, plutôt pas d'accord ;

- Les cotations 3 et 4 correspondent à un accord avec la proposition : 4, tout à fait d'accord ; 3, plutôt d'accord ;

- la colonne NSP correspond au nombre d'absence de réponse ou de « je ne sais pas » ; NSP, ne se prononce pas.

- afin de distinguer le positionnement des élèves, le choix a été fait de ne pas utiliser les indicateurs classiques des statistiques descriptives d'autant que les effectifs des différents échantillons sont faibles (le nombre d'élèves tourne autour d'une vingtaine par classe) ;

- trois indicateurs utilisés pour analyser les réponses des élèves sont les suivants : le nombre de réponses « d'accord », « de désaccord », et la cotation la plus fréquente.

- afin de préserver l'aspect du microcosme constitué par la classe à un moment donné, les items proposés au professeur et aux élèves sont symétriques : ils sont autant que possible rédigés de la même façon ; sinon, leur rédaction fait référence à un vécu commun. Par exemple, on demande au professeur s'il estime que « le temps était suffisant », et aux élèves « s'ils ont eu assez de temps »

Questionnaire professeur		Questionnaire élèves				
Item	Cotation	Cotation la plus fréquente	Nb de réponses 1 ou 2	Nb de réponses 3 ou 4	NSP	
Le temps était suffisant pour que les élèves finissent	4	4	9	14	0	J'ai eu assez de temps pour finir
L'énoncé s'inspire de ce qui a été fait en classe	4	4	4	19	0	J'ai reconnu ce qu'on avait fait en classe
C'est un contrôle où il faut refaire ce qu'on a fait en classe	3	2	10	11	2	C'est un contrôle où il faut refaire ce qu'on a fait en classe
La réponse était dans le cahier	4	1	12	11	0	
C'est une question facile quand on a bien révisé	3	2	11	12	0	

Analyse des réponses :

- alors que le professeur estime que les élèves avaient assez de temps pour finir, seulement 14 élèves sur 21 ont la même appréciation. On peut noter également que la réponse la plus fréquente donnée par les élèves correspond à celle du professeur, mais que des indicateurs classiques de statistiques descriptives (moyenne, médiane, écart-type) n'aurait pas mis en évidence aussi nettement l'existence des 7 élèves en désaccord avec leur professeur. En ce sens, la méthodologie retenue donne des éléments d'analyse variés et faciles dans leur exploitation.

- L'item « il faut refaire ce qui a été fait en classe » donne différentes indications :
- Le professeur estime qu'il ne demande pas exactement ce qui a été fait en classe en répondant 3 (pas tout à fait d'accord) : cela questionne sur ses intentions évaluatives en sortant du curriculum réel, et peut-être également sur la façon dont il respecte le contrat didactique du moment ;
- La réponse la plus fréquente des élèves est 2 (pas d'accord) ;
- Les réponses des élèves se partagent et surtout 11 d'entre eux ne reconnaissent pas ce qui a été fait en classe et donc l'abordent sans doute avec toute l'appréhension imaginable face à des questions nouvelles.
- L'item « la réponse était dans le cahier » prête à réflexions : le professeur est « tout à fait d'accord » (cotation 4) ; en revanche, 11 élèves ne s'en sont pas rendu compte. Alors que le professeur pensait avoir posé une question « facile » puisque la réponse était dans le cahier, 11 élèves se sont trouvés dans l'incapacité d'en profiter.

Résultats, bilan et perspectives

La méthodologie retenue permet de conclure à l'existence d'un écart systématique entre les réponses du professeur et celles des élèves : en comparant les scores globaux, aucun élève n'a obtenu le même score que son professeur. Cela signifie donc qu'aucun élève n'a totalement ni exactement perçu les intentions évaluatives du professeur, ce qui signifie qu'il existe un écart. Dans aucune des études de cas, deux élèves ont obtenu le même score global de la même façon, ce qui confirme l'existence de la conscience évaluative individuelle.

Cette réalité a une influence sur la performance des élèves et donc sur le jugement évaluatif que le professeur va porter sur chacun d'eux. Or, de la recherche précédente, il ressort que trois paramètres relevant de la didactique générale permettent au professeur d'analyser l'énoncé qu'il élabore : dans sa globalité est-il conforme au curriculum réel, au contrat didactique du moment notamment à travers les tâches demandées ; dans le détail, chaque question est-elle conforme au curriculum réel, au contrat didactique y compris dans le choix des tâches demandées. **Cette analyse, seul le professeur qui a assuré les apprentissages est en mesure de la mener, c'est-à-dire le professeur de la classe.**

Les paramètres retenus ici ne sont pas exclusifs, mais en revanche, ils ont l'intérêt de conforter la place de l'énoncé de l'épreuve comme étant le centre du processus évaluatif interactif. D'autres paramètres peuvent contribuer à analyser un énoncé, par exemple ceux relevant d'une classification des tâches demandées : ils donnent une autre description, « neutre » de l'énoncé, alors que la recherche présentée ci-dessus vise à montrer l'importance du caractère contextuel d'une évaluation interne.

Ces résultats sont obtenus dans les disciplines scientifiques enseignées au collège, parce la forme des énoncés et des réponses attendues est assez semblable dans les quatre disciplines concernées : un travail complémentaire équivalerait mériterait d'être mené au niveau de l'école élémentaire et des lycées (LP et LEGT). Dans les lycées, une telle investigation si les mêmes conclusions sont vérifiées, apporterait de nouveaux éléments dans les choix d'orientation faits par les élèves et les décisions prises par les conseils de classe, dans la mesure où ils interrogent les conditions dans lesquelles la performance est réalisée. Une autre extension serait à envisager dans les disciplines d'enseignement où les énoncés peuvent avoir des formes différentes : les disciplines relevant des

sciences humaines (LV, histoire-géographie-éducation civique, français) et d'expression-création (EPS, arts plastiques, éducation musicale) : il serait nécessaire de reconsidérer les items utilisés pour connaître les intentions évaluatives du professeur et la conscience évaluative des élèves, par exemple en français, quand l'activité porte sur une rédaction à produire sur un thème proposé.

Enfin, l'appréhension que les élèves ont des intentions évaluatives de leur professeur a été exploré ici dans le cadre didactique : son étude dans le cadre sociologique (la dualité élève/enfant-adolescent ; l'influence du groupe classe sur la conscience évaluative de chaque élève selon son statut dans la classe) ou psychologique (la relation spécifique du professeur avec chaque élève, l'importance donnée à l'école dans la famille) serait de nature à mieux comprendre le comportement d'un élève engagé dans une situation d'évaluation formelle et donc sa conscience évaluative.

Enfin, l'influence de cette compréhension des intentions évaluatives du professeur par les élèves sur leurs performances, incite à s'interroger sur l'existence d'une réalité comparable dans les phases d'apprentissage. En transposant ce qui a été exploré ici dans les moments scolaires consacrés à l'évaluation, aux moments durant lesquels les apprentissages sont menés, la question de la réceptivité des élèves et spécifiquement de chaque élève pourrait être posée : l'hypothèse de l'influence de cette réceptivité (liée à ce que pourrait être une « conscience didactique ») mériterait sans doute d'être explorée dans la mesure où elle pourrait expliquer des difficultés momentanées, plus durables voire même des refus d'apprentissage de certains élèves.

Bibliographie

Brousseau G. (1997). Le contrat didactique : le milieu. *Recherches en didactique des mathématiques*, 9 (9.3), 309-336.

Grawitz, M. (2001). *Méthodes des sciences sociales* (11e éd.). Paris : Dalloz.

Issaieva E. , Pini G. & Crahay M. (2012). Positionnements des enseignants et des élèves du primaire face à l'évaluation : une convergence existe-t-elle ? *Revue française de pédagogie* 176(3), 5-26.

Jorro, A. (2000). *L'enseignant et l'évaluation : des gestes évaluatifs en question*. Bruxelles : De Boeck supérieur.

Figari, G & Mottier Lopez, L. & ADMEE Europe (Eds). (2006). *Recherche sur l'évaluation en éducation : problématiques, méthodologies et épistémologie : 20 ans autour de l'ADMEE-Europe*. Paris : L'Harmattan.

Merle, P. (2018). *Les pratiques d'évaluation scolaire*. Paris : PUF.

Perrennoud, P. (1993). In Houssaye, J. (dir.) *La pédagogie : une encyclopédie pour aujourd'hui*, Paris, ESF. pp. 61-76.

Raulin, D. & Lebeaume, J. (2017). L'élève acteur de l'évaluation. In A. Jorro & N. Droyer (dir). *L'évaluation, levier pour l'enseignement et la formation*. Bruxelles : De Boeck supérieur.

Reuter, Y. (2008). *Dictionnaire des concepts fondamentaux des didactiques* (2e éd.) Bruxelles : De Boeck.

Rey, B. (2012). *Les compétences à l'école : apprentissage et évaluation*. Bruxelles : De Boeck.