

Suisse : transformations des énoncés d'évaluation dans les épreuves écrites de philosophie dans l'enseignement secondaire postobligatoire genevois, 1980-2010

Léo Freuler, professeurs de philosophie du canton de Genève

Simone Delorenzi, professeurs de philosophie du canton de Genève

Nous publions ci-dessous, de façon simplifiée, une "Étude de docimologie diachronique" réalisées par deux collègues suisses sur l'évaluation en philosophie dans le canton de Genève

Première Partie. L'objet et la méthode

A) Délimitation de l'objet étudié

1. Le travail présenté ici concerne le domaine des transformations du travail de l'enseignant, non pas dans une perspective comparée, mais historique, ou diachronique. Notre travail porte sur l'évaluation, et il relève par conséquent de ce que, depuis Piéron, on appelle la docimologie (Martin, 2002). La discipline à travers laquelle nous allons étudier certains aspects de l'évaluation est la philosophie, qui est notre discipline d'enseignement. Cependant, il est à noter que nous ne présentons pas ici un travail sur l'évaluation en philosophie. Nous traiterons les épreuves de philosophie comme un exemple permettant d'étudier des aspects techniques de l'évaluation en général. Aussi notre travail ne relève-t-il pas de la didactique de la philosophie, mais de la profession enseignante en général. Parce qu'il se contente d'étudier des énoncés d'épreuves de philosophie, notre travail ne sera donc pas transdisciplinaire. Mais parce qu'il ne s'intéresse qu'aux caractéristiques générales de l'évaluation, il pourrait très bien devenir le point de départ d'une étude transdisciplinaire.

2. Aspects docimologiques couverts par le travail

Dans la question de l'évaluation, on peut distinguer quatre éléments (Freuler, 2010) :

- La personne qui évalue (sujet évaluant), c'est-à-dire l'enseignant (a);
- La personne qui est évaluée (sujet évalué), c'est-à-dire l'élève (b);
- Le contenu de l'évaluation, ou l'objet par lequel on évalue (objet évaluant), c'est-à-dire l'énoncé de l'épreuve (c);
- L'objet qui est évalué (objet évalué), c'est-à-dire l'épreuve (la production de l'élève) (d).

On peut concevoir au moins six relations (I-VI) possibles entre ces éléments, et les représenter sous la forme de ce qu'on peut appeler le carré docimologique.

Ce carré pourrait faire l'objet de nombreux développements théoriques, lesquels, cependant, relèveraient d'un travail différent. Dans le cadre qui est le nôtre, nous nous limiterons à certains aspects de ce carré.

Notre objet principal, c'est l'objet évaluant (a), c'est-à-dire les énoncés d'épreuve. Depuis les célèbres expériences de multicorrection des années 1930, qui ont ébranlé, toutes disciplines confondues, le mythe de l'objectivité des corrections¹, la recherche docimologique s'intéresse beaucoup, sinon principalement, aux problèmes de l'erreur et de la subjectivité de l'évaluation (élément a). Notre but n'est donc pas de contribuer à cette partie de la docimologie, mais à une autre qui étudierait spécifiquement les énoncés d'évaluation, considérés en eux-mêmes.

L'étude de cet objet principal nous amènera cependant à considérer certains aspects des relations I et IV, par exemple pour éclairer la manière dont un enseignant, à travers son énoncé, choisit de s'adresser à ou de guider son élève, ou pour tirer des enseignements sur notre propre pratique dans la formulation d'énoncés.

B) Méthode et objectifs

Concernant la méthode de notre recherche, notre travail se servira de deux sortes de méthodes :

- la méthode empirique ou inductive pour l'observation de régularités et d'irrégularités ;
- la méthode abductive, pour chercher des hypothèses qui permettraient de rendre compte de ce qui a été observé.

Nous ne procéderons donc pas de manière hypothético-déductive, nous ne formulerons pas des hypothèses dans le but de les vérifier. Nous commencerons par étudier des faits, c'est-à-dire le corpus de notre recherche, et nous chercherons des hypothèses susceptibles d'en rendre compte. En d'autres termes, nous ne nous proposons pas, dans ce travail, de vérifier des hypothèses, mais de formuler des hypothèses, à partir de l'exploration du corpus d'énoncés à notre disposition.

D'autre part, nous allons combiner ces deux méthodes avec le point de vue diachronique propre notre étude.

En un premier temps, nous allons procéder à des études purement empiriques des transformations qu'on peut observer dans les énoncés d'épreuves. C'est l'aspect diachronique de notre travail. Nous rechercherons donc l'apparition de nouveautés et d'innovation et nous les comparerons à l'état avant leur apparition, ce qui nous permettra d'observer des continuités, des évolutions ou des ruptures.

En un deuxième temps, nous allons chercher des hypothèses qui permettraient de rendre compte de ces transformations. C'est l'aspect abductif de notre travail. Ces deux moments constituent aussi les objectifs de notre travail.

C) Matériaux

Concernant le matériau de notre étude, il est constitué de documents officiels, en l'espèce d'énoncés d'épreuves écrites de philosophie, provenant d'établissements de l'enseignement secondaire postobligatoire du canton de Genève.

Les énoncés que nous avons pu consulter faisaient toujours partie de lots, ou de collections. En tout, nous avons eu accès à environ six cents cinquante épreuves. Nous avons pu accéder à des énoncés d'épreuves couvrant la période de l'année scolaire 1980-1981 jusqu'à l'année scolaire 2009-2010. Notre étude comportant un important volet diachronique, nous avons systématiquement écarté les énoncés dépourvus de date.

Enfin, le matériau comporte des épreuves destinées à des 3e années et des 4e années, ainsi que des épreuves de tous types (regroupées, semestrielles, de maturité, récapitulatives, d'option complémentaire, d'option spécifique, trimestrielles, intermédiaires).

Deuxième Partie : analyse et hypothèses

A) Premier résultat général de la recherche

En parcourant de manière empirique des énoncés d'épreuve de philosophie, trois points se sont progressivement imposés à notre attention, en raison à la fois de l'importance particulière que ces éléments revêtent dans les épreuves, ainsi que du caractère frappant de leurs évolutions respectives. Nous avons décidé de nous concentrer sur ces trois points :

1. La nature et la mention de consignes.

Nous avons été frappés par le fait que, au début de notre limite chronologique, on ne trouve aucune, ou très peu de consignes, et qu'elles ne commencent à se multiplier que vers 2000, pour devenir des parties intégrantes des énoncés.

2. La nature et la mention des critères d'évaluation.

Là aussi, l'intérêt particulier pour cet aspect des énoncés est lié au fait que, au début de la limite chronologique de notre corpus, on ne trouve presque aucune mention de critères d'évaluation, et qu'il faut attendre les années 2000 pour les voir apparaître et se multiplier.

3. Le type de question.

Nous faisons encore le même constat pour ce troisième aspect des énoncés. Progressivement, les questions posées se compliquent et se fragmentent. D'une petite quantité de questions laissant un spectre très large au candidat (ce qu'on pourrait appeler des macroquestions), on passe à des questions de plus en plus pointues et resserrées, ciblant des objets de plus en plus précis et exclusifs (ce qu'on pourrait appeler des microquestions).

B) État de la bibliographie

Nous allons procéder à une étude plus approfondie de ces trois points.

Signalons au préalable que nous ne pourrions pas toujours nous appuyer, dans cette démarche, sur une bibliographie spécifique, notamment en ce qui concerne les deux premiers points. Ainsi, on a certes développé la réflexion sur les consignes, mais plus particulièrement dans la direction de la réception des consignes par les élèves, comme on peut le voir chez Anonyme(s) (1997), Raynal &

Rieunier (1997), ou encore Zakhartchouk (1999). En revanche, nous n'avons rien trouvé de systématique concernant l'aspect qui nous intéresse, c'est-à-dire la mention des consignes et leur typologie.

Quant à la mention de critères d'évaluation sur les énoncés, nous n'avons, là non plus, rien trouvé de systématique concernant l'évolution et la typologie de ces critères. On trouve plus de choses, par contre, sur l'évaluation quantitative, mais de telle manière que, là non plus, elles ne concernent pas l'objet de notre étude. Car si on trouve énormément d'exemples de barèmes et de grilles d'évaluation, ou encore des guides pratiques de l'évaluation, p. ex. chez Morissette (1996), on cherche en vain des études plus théoriques sur l'évolution et la typologie des critères mentionnés sur les énoncés. Là aussi, nous avons dû explorer ce qui pourrait être une sorte de terrain vierge, et construire empiriquement un appareil conceptuel susceptible de l'éclairer.

C'est dans la typologie des modèles de l'évaluation, qui concerne notre troisième point, qu'on trouve la littérature la plus abondante.

C) Etude particulière des trois points

1. Nature et mention des consignes

a) Constats

--> État de la question dans les années 1980 et 1990

D'une manière générale, on peut constater, entre les années 1980 et 2010, une tendance assez claire à la multiplication des consignes. Avant les années 2000, on ne trouve presque que des énoncés qui laissent une liberté maximale aux élèves. Les questions sont extrêmement courtes, et presque sans consignes particulières. Souvent les énoncés ne remplissent même pas une page entière. Parfois, on n'y trouve aucune consigne du tout, parfois des consignes minimalistes, concernant par exemple la "brièveté" et la "concision" des réponses.

Au début des années 1990, la majorité des énoncés sont toujours dépourvus de consignes. Vers la fin des années 1990, on ne constate guère d'évolution notable dans cette question. Cependant, on en voit apparaître qui, de toute évidence, reflètent des préoccupations nouvelles) : "Traitez **l'un** des sujets proposés ci-dessous et numérotés de 1 à 5. Les références aux pages du livre de Quillot pourront servir de bases argumentatives pour votre débat. Elles ne sont évidemment pas limitatives mais suggestives ! Il ne s'agit pas non plus de les recopier servilement !".

Ces consignes ne manquent pas d'étonner. Dans la première, l'examineur met en gras l'expression l'un, comme si la compréhension de cette formule n'allait plus de soi pour certains élèves. Il a également pris soin de signaler la numérotation des questions, comme si, là aussi, certains élèves pourraient se tromper sur le nombre de questions à traiter. Quant à la deuxième consigne, elle a

manifestement pour but de mettre en garde contre une certaine passivité, comme si l'examineur avait constaté une perte dans la capacité d'autonomie ou de création libre de ses élèves.

--> Etat de la question en 2005-2006

Faisons maintenant un bond en avant, et consultons quelques énoncés des années 2005-2006. Dans certains d'entre eux, on voit apparaître des encadrés intitulés "Consignes". En voici un exemple tiré d'une épreuve du 17 janvier 2006 :

Consignes :

- Soignez la présentation !
- Ecrivez lisiblement à l'encre
- Signalez les citations par des guillemets ("...").

Dans la même épreuve, ces recommandations "particulières à la discipline" sont précédées d'une série de "Recommandations générales" :

- Sur la première page des feuilles d'épreuves, veuillez vous limiter aux informations administratives, à savoir votre nom, la date le nom du maître de la discipline, et commencez l'épreuve proprement dite à la page suivante.
- Notez ensuite votre nom en haut de chaque page et numérotez-le.
- Notez sur la première page l'heure de reddition de votre travail.
- N'oubliez pas de rendre l'énoncé avec votre travail à la fin de l'épreuve.

Toujours dans la même épreuve, on donne également des consignes sur la manière de traiter les questions :

Vous avez 12 questions

Il faut donner des réponses concises mais complètes. Développer en fonction de la pondération des questions. Deux trois phrases suffisent pour les questions courtes. Veiller à bien enchaîner les arguments lorsqu'un raisonnement est nécessaire, afin d'amener le lecteur à une conclusion claire. Les explications doivent être les plus explicites possibles".

Cette tendance aux rubriques entières consacrées aux consignes se constate dans plusieurs établissements, et il arrive encore qu'on les regroupe sous le terme de "Rappels", comme si les élèves avaient de la peine à les mémoriser.

Au lieu d'être regroupées, il arrive aussi que, en 2006 toujours, les consignes soient dispersées sur l'ensemble de l'énoncé, jalonnant la feuille d'épreuve du début jusqu'à la fin. Dans certains d'entre eux, on va même jusqu'à expliquer à l'élève qu'il faut tourner la page : "Tournez s.v.p."

b) Tendance générale et nuances

Ces exemples de 2006 montrent que nous sommes loin des années 1980 et de son extrême parcimonie en matière de consignes. A partir de 2000 environ, on peut clairement observer une multiplication des consignes. Cependant il faut prendre soin de nuancer ce constat, et cela sous deux rapports au moins :

S'il s'agit d'une multiplication des consignes, il ne s'agit ni d'une généralisation, et encore moins d'une universalisation. En effet, le recours à la consigne est très variable d'un enseignant à un autre. Certains enseignants y recourent très peu, et demeurent constants dans leur pratique, y compris en 2006. D'autres y recourent beaucoup et demeurent également constants dans leur pratique. On peut noter également des pratiques concertées dans certains établissements.

Par ailleurs, si les consignes se multiplient avec le temps, on en trouve déjà dans les années 1980, bien que le phénomène soit beaucoup plus rare, pour ne pas dire exceptionnel. Ainsi, dans une épreuve du Collège Voltaire, on peut trouver des "remarques" en haut de la feuille et en bas de la feuille, les premières destinées aux maîtres, les deuxièmes destinées aux élèves: "Vous commenterez, illustrerez d'exemples, discuterez chaque point de la pensée choisie et vous exposerez de façon synthétique, dans la conclusion de votre dissertation, votre solution au problème posé".

Dans une épreuve de l'année scolaire 1989-1990, on trouve un certain nombre de consignes qui concernent le nombre de questions et la longueur de la réponse: "en quelques mots ou en quelques lignes", "en 1/2 page environ", "en 2 pages environ".

Au vu de ces exceptions et de ces variations d'un enseignant l'autre, tenons-nous donc à notre formulation : concernant la mention de consignes sur les énoncés d'épreuves, nous constatons une multiplication de ces consignes à partir de l'an 2000 environ.

c) Multiplication des consignes et typologie des consignes

Nous pouvons encore affiner notre observation en essayant de déterminer si cette multiplication concerne tous les types de consignes, ou seulement certains d'entre elles. Il est clair, en effet, que les consignes ne sont pas toutes de même nature. Voici une typologie de consignes figurant sur les énoncés consultés, résultat d'une analyse purement empirique:

- Consignes quantitatives : nombre de mots, de lignes, de pages ; nombre de questions.
- Consignes qualitatives : concernant la qualité de la démarche scientifique (donner des arguments, être clair, etc.) ; concernant le contenu : pas de "blabla", chercher dans telle direction plutôt que telle autre, consulter tel texte, etc. ; concernant les règles de la citation : mettre des guillemets, etc. ; concernant la qualité de la langue : soigner l'orthographe, la syntaxe, etc.
- Consignes matérielles : concernant la feuille (rendre la feuille à la fin de l'exposé, noter l'heure de la reddition) ; concernant l'écriture : écrire lisiblement, écrire à l'encre ("Remarque : Votre travail devra être rédigé à l'encre") ; concernant les indications devant figurer sur la page : noter

le nom, la date, le nom du maître ; concernant la mise en page : numérotation des pages, lignage des marges ; concernant la manipulation des pages : "tournez, s.v.p".

Essayons d'appliquer cette typologie au cadre chronologique qui est le nôtre.

Si on observe les quelques très rares consignes qui apparaissent dans les années 1980 et les années 1990, on peut constater qu'elles sont toutes, et très clairement, de nature qualitative. Et elles ne concernent que la qualité de la démarche scientifique et le contenu. En revanche, on ne trouve, pendant cette période, pas de consignes concernant l'expression écrite, c'est-à-dire concernant les règles de la citation et la qualité de la langue. Certes, nous n'avons pas pu vérifier cela sur l'ensemble de tous les énoncés délivrés dans tous les établissements du canton. Cependant on voit clairement une différence entre les années antérieures à 2000 et l'état de la question dans une année comme 2006. En 2006, presque tous les énoncés comportent des consignes concernant l'expression écrite. La nouveauté des années 2000, par conséquent, c'est l'apparition massive de trois sortes de consignes : consignes qualitatives concernant l'expression écrite ; consignes quantitatives ; consignes matérielles.

d) Hypothèses explicatives

Tel est l'état descriptif de la question des consignes dans le corpus d'évaluations que nous avons examiné. Essayons maintenant de formuler quelques hypothèses explicatives sur la tendance majeure constatée, c'est-à-dire sur la multiplication des consignes concernant l'expression écrite, des consignes quantitatives et des consignes matérielles.

- Causes endogènes et causes exogènes

Une première hypothèse consiste à dire, évidemment, que cette tendance est imputable à une baisse objective des compétences correspondantes des élèves. Avec le temps, et plus particulièrement à partir de 2000, certains élèves manifesteraient une perte de compétence dans l'expression écrite, dans l'usage du matériel d'épreuve, ou encore dans la maîtrise du nombre de questions et de la longueur de la réponse.

Si cette hypothèse était confirmée, nous devrions chercher ensuite les causes de cette baisse de compétence. La première possibilité est qu'elle est due à des causes exogènes, la deuxième qu'elle est due à des causes endogènes, la troisième enfin qu'elle est due aux deux sortes de causes deux en même temps. Si elle est due à des causes exogènes, elle dépend de facteurs extérieurs à l'école, par exemple une baisse croissante des compétences de mémorisation et de concentration, qui oblige à tout fixer par écrit, ou encore à l'invasion des images au détriment de l'écrit, et donc des compétences lectorales. Dans ce cas, l'école n'aurait pas eu assez de pouvoir pour contrebalancer ces causalités externes, ou au contraire en aurait été assez complice pour ne pas avoir la volonté d'en corriger les effets.

Si par contre la ou les causes sont endogènes, elle pourrait être due à des attitudes des enseignants eux-mêmes. Or qu'est-ce qui aurait pu pousser une partie des enseignants à introduire des consignes qualitatives concernant l'expression écrite, ainsi que des consignes quantitatives et matérielles ? A ce sujet, nous pouvons formuler au moins quatre hypothèses, ou hypothèses endogènes.

- Une première hypothèse endogène ferait le lien avec d'éventuelles causes exogènes et consisterait à dire qu'il s'agit d'une réponse des enseignants à une baisse objective des compétences correspondantes dues à des causes exogènes. Dans ce cas, les enseignants auraient noté des évolutions négatives qui ne sont pas dues à l'école elle-même, et auraient décidé de réagir en rappelant quelques consignes élémentaires que leurs prédécesseurs n'estimaient même pas dignes d'être énoncés, parce qu'elles étaient évidentes, ou constituaient des acquis élémentaires. Dans cette hypothèse, on serait dans une logique réactive.

- Une deuxième hypothèse endogène consisterait à dire que certains enseignants, voire une nouvelle génération d'enseignants, sont devenus plus méticuleux parce que plus directifs, c'est-à-dire plus enclins à vouloir orienter, voire contrôler les élèves pour les pousser dans une certaine direction. Dans leur esprit, peut-être, l'élève n'est pas d'abord et principalement un être qu'il faut éduquer à l'autonomie en le laissant expérimenter cette autonomie, mais qu'il faut guider d'une certaine manière pour le conduire vers certains objectifs. C'est pour s'en assurer qu'on multiplie les consignes. Dans cette hypothèse, on serait donc dans une logique directive.

- Une troisième hypothèse consisterait à dire que certains enseignants, voire une nouvelle génération d'enseignants, sont devenus plus angoissés. Ils ont le souci de la réussite de leurs élèves, et par conséquent craignent leur propre échec. Aussi décident-ils de multiplier les consignes, moins pour contraindre les élèves dans une certaine direction, que pour diminuer le risque d'échec, pour créer un maximum de sécurité et se garantir un maximum de l'échec. Dans cette hypothèse, on serait donc dans une logique sécuritaire. L'intention serait fondamentalement bonne, ou bienveillante, mais on court le risque d'infantiliser les élèves et de les priver d'une certaine autonomie.

- Une quatrième hypothèse pourrait être fondée sur des considérations beaucoup plus pragmatiques. Il est certain que, dans les années 1980 et 1990, les enseignants n'étaient pas moins sensibles à la qualité de l'expression écrite que leurs successeurs dans les années 2000. Pourquoi donc, après l'an 2000, a-t-on pris soin de mettre ce genre de consigne par écrit ?

Scripta manent, dit l'expression. En d'autres termes, la formulation écrite constitue une référence permanente à laquelle on peut se reporter en cas de doute, non seulement pour se rappeler les consignes au moment du travail, mais également au moment de son évaluation. L'enseignant a clairement fixé les exigences, et l'élève ne pourra pas prétendre, après coup, que les choses n'ont pas été dites et précisées.

On voit que l'hypothèse concerne donc l'idée que, avec le temps, les élèves ont pris l'habitude de contester les corrections, notamment en prétendant qu'on ne peut pas leur imputer certaines erreurs

parce que les exigences correspondantes n'ont pas été formulées. Aussi certains enseignants, pour se prémunir contre des contestations, auraient-ils pris l'habitude d'indiquer par écrit un maximum de règles. Pour la même raison, le législateur met-il par écrit toutes ses lois et les rend publiques, afin que tout le monde soit averti : nul n'est censé ignorer la loi, c'est-à-dire tenter de s'innocenter après coup en invoquant l'ignorance de la loi. Cette hypothèse, on le voit, s'inscrit également dans une logique sécuritaire, tout comme la précédente. Contrairement à la précédente, cependant, son but n'est pas de prévenir l'échec de l'élève, mais la contestation de l'élève.

Dans ce sens, parce qu'elle concerne la contestation de l'élève, on pourrait également la lier à la question de la perte d'autorité des enseignants, au fait que, après les coups portés contre le modèle traditionnel de ce que Robbes (2006) appelle l'"autorité autoritaire", on ait assisté à un déclin progressif de l'autorité, jusqu'à l'absence revendiquée d'autorité, que Robbes qualifie d'"autorité évacuée". Sous le régime de l'autorité autoritaire, rien n'était contestable, tandis que sous celui de l'autorité évacuée, tout le devient. Or la multiplication des contestations et, par réaction, celles de certaines consignes, ne serait-elle pas précisément un effet de ce glissement vers l'autorité évacuée, vers l'absence ou le refus d'autorité, parfois revendiqué par certains professeurs eux-mêmes ? Dans ce cas, on aurait déterminé une cause exogène et première de la multiplication de certains types de consignes. On conteste l'autorité du professeur, et par réaction, certains professeurs multiplient les garde-fous pour se protéger contre ces contestations.

2. Nature et mention des critères d'évaluation

Passons maintenant à la question de la mention des critères d'évaluation. Rappelons qu'il ne s'agit pas d'étudier ici ces critères en eux-mêmes, mais seulement la question de leur apparition ou mention sur les feuilles d'énoncés.

L'un des points frappants qui ressort de l'étude de ce point est qu'on y retrouve à peu près les mêmes aspects fondamentaux que dans la question des consignes. Pour s'en apercevoir, inversons l'ordre de recherche et commençons par la typologie des critères, pour ensuite examiner l'évolution de leur mention.

a) Typologie des critères d'évaluation

Les critères d'évaluation qui apparaissent dans notre corpus d'énoncés peuvent être classés en deux types fondamentaux, dont chacun comporte à son tour deux sous-types :

Critères qualitatifs : concernant l'expression (orthographe, syntaxe, style) ; concernant le contenu scientifique (richesse et pertinence du contenu, cohérence).

Et critères quantitatifs : points pour le contenu (barèmes) ; points pour l'orthographe.

--> État de la question dans les années 1980 et 1990

Examinons maintenant l'évolution de la mention de ces critères d'évaluation dans les énoncés de notre période. Le résultat général de cet examen est qu'on peut constater ici quelque chose d'analogue à l'évolution de la mention des consignes. Dans les années 1980 et 1990, en effet, on ne trouve presque aucun énoncé qui fasse explicitement mention des critères d'évaluation. Dans les années 1980, on peut trouver au moins deux énoncés qui mentionnent des critères quantitatifs, de telle manière cependant qu'ils ne s'adressent pas aux élèves, mais constituent des sortes de notes internes à l'usage du correcteur :

[En haut de la feuille:] Remarques particulières : (Les maîtres doivent préciser, le cas échéant, s'il faut traiter toutes les questions proposées ou s'il s'agit de sujets à choix et indiquer, si possible, l'importance de chaque questions dans l'évaluation générale).

[En bas de la feuille, ajouté à la main :] 6 points par question. Barème linéaire.

Nous avons également trouvé deux énoncés, émanant du même établissement, qui donnent des indications quantitatives à l'intention des élèves. Il ne s'agit pas, cependant, de barèmes ou de points, mais de la longueur des réponses ou le nombre de mots :

- Question 1 : "Expliquez, commentez cette phrase (maximum 1/2 page)".
- Question 2 : "maximum une 1/2 page".
- Question 4 : "maximum une 1/2 page".

Traitez les questions 1 et 2 en 120 à 150 mots, la question 3 en 250 à 300 mots; les réponses ne doivent comprendre aucune citation.

Dans la décennie suivante, un seul énoncé évoque la question de l'évaluation, cette fois de manière plus complète, et à l'intention des élèves. Dans cette exception, cependant, on ne traite que de critères qualitatifs (richesse et pertinence du contenu, cohérence interne, style et syntaxe, orthographe et accords). On ne donne aucun nombre de points, on ne voit apparaître aucun barème, ni aucune formule permettant de calculer la note. A la fin, on conclut sur une note humoristique en donnant la consigne de "dormir la nuit qui précède l'examen".

--> Évolution de la question après 2000

Franchissons maintenant le cap de l'an 2000. Une première chose qu'on peut alors observer, c'est une réduction des critères d'évaluation aux critères quantitatifs. Ce n'est pas que les critères qualitatifs disparaissent des énoncés, mais c'est qu'ils apparaissent plutôt du côté des consignes, et ne sont pas spécifiquement mentionnées comme des critères d'évaluation. C'est comme si, en d'autres termes, il y avait une sorte de divorce entre les critères qualitatifs et les critères quantitatifs.

Concentrons-nous donc sur les critères quantitatifs. Dans notre corpus, les premières mentions d'un critère d'évaluation quantitatif apparaissent en 2001. Dans un énoncé du CEC Emilie-Gourd, on mentionne des points et on donne une formule permettant de calculer la note : "Evaluation : nombre de points obtenus \times 5/20 + 1 = note".

Faisons maintenant un bond vers les années 2005 et 2006. En 2005, aucun énoncé de notre corpus ne comporte la mention de critères quantitatifs d'évaluation. En 2006, en revanche, on voit apparaître plusieurs énoncés mentionnant des critères quantitatifs, bien qu'ils restent clairement minoritaires. Les types de critères quantitatifs qui font leur apparition sont les suivants : points pour chaque question ; mention du total des points : formule permettant de calculer la note ; points pour l'orthographe.

b) Tendance générale et nuances

Ces exemples montrent assez clairement que, comme dans le cas des consignes, quelque chose a changé dans les années 2000. Mais comme dans le cas des consignes, on ne saurait parler ici de la généralisation, voire de l'universalisation d'une pratique. On ne peut même pas parler d'une pratique majoritaire, car une telle conclusion nécessiterait des études statistiques et un corpus de recherche dont nous ne disposons pas. Tout ce qu'on peut objectivement constater, c'est une multiplication des critères d'évaluation quantitatifs, dont la nature et l'intensité peuvent varier d'un établissement à un autre, et même d'un enseignant à un autre (cf. par exemple quatre épreuves de quatre enseignants différents du même établissement).

c) Hypothèses explicatives

Si nous passons maintenant aux hypothèses qui pourraient rendre compte de cette tendance à la multiplication de critères quantitatifs d'évaluation, nous pouvons, là encore, trouver des analogies remarquables avec la question des consignes. Nous pouvons reprendre, en effet, la plupart des hypothèses mentionnées dans le cas de la tendance générale dans la question des questions, et les appliquer à celle de la mention des critères de l'évaluation. Nous pouvons donc nous demander si l'irruption de critères quantitatifs est due à des causes exogènes ou endogènes, et si, dans le deuxième cas, on peut les expliquer par l'une de ces quatre hypothèses qualifiées d'endogènes. C'est le cas, du moins, des trois dernières hypothèses endogènes, car la première¹, qui consistait à expliquer la multiplication des consignes par une réaction des enseignants à une perte objective de certaines compétences des élèves dues à des causes exogènes, ne s'applique manifestement pas à la question des critères quantitatifs. En revanche, nous pouvons nous demander si l'apparition de critères quantitatifs correspond, du côté des enseignants, à (2) un besoin de mieux contrôler les élèves, à (3) un besoin sécuritaire de prévenir l'échec de l'élève (et donc également de l'enseignant lui-même, parce qu'il a échoué à bien préparer son élève), à (4) un besoin de fixer des critères pour prévenir des contestations des élèves.

Dans ce dernier cas, la mention d'une évaluation quantitative comporte encore une dimension supplémentaire et peut-être spécifique qui n'apparaît pas dans la question des consignes. En effet, une telle mention ne peut pas seulement avoir pour but de prévenir d'éventuelles contestations parce qu'elle est faite, donc annoncée publiquement et par écrit (nul n'est censé ignorer le critère). C'est la raison de la mention qui nous est apparue dans le cas des consignes. Elle peut également viser le même but parce qu'elle mentionne quelque chose de quantitatif. Ainsi un barème, cela peut donner l'impression d'être objectif comme un théorème de mathématique ou un axiome de géométrie et, par conséquent, cela peut donner l'impression d'être incontestable. Dans ce cas, le but de l'enseignant serait donc d'éviter la contestation en évitant l'arbitraire de la notation, de se servir de critères quantitatifs pour faire croire à l'objectivité de l'évaluation.

Dans le cas de la mention des critères d'évaluation, cette quatrième hypothèse touche donc également à la question de l'"objectivité" de l'évaluation, et elle a des liens, de ce fait, avec la problématique lancée par les fameuses expériences de multicorrection des années 1930. Cette question, cependant, dépasse le cadre de notre travail et devrait faire l'objet d'un autre2.

D'autre part, la volonté de prévenir la contestation par ce biais pourrait également être liée, comme dans la question des consignes, à la perte d'autorité des enseignants depuis les années 1960, au basculement d'une conception de l'autorité de soumission, ou autoritaire, à une autorité refusée ou "évacuée", y compris parfois par les enseignants eux-mêmes. Plus leur autorité a été remise en question, plus certains enseignants auraient décidé, par réaction, de multiplier certains garde-fous, en l'espèce des critères quantitatifs susceptibles de mettre un terme à toute espèce de contestation.

3. Nature des questions

Passons maintenant au dernier de nos trois points, celui qui concerne la nature des énoncés. D'une manière générale, on peut constater, entre les années 1980 et 2010, une tendance à la multiplication des types d'exercices proposés, reflétant dans certains cas une tendance à l'élémentarisation des connaissances et, surtout, des compétences requises pour effectuer la tâche demandée.

Nous retrouvons donc, au niveau même des types d'énoncés, la même tendance inflationniste que nous avons pu observer au sujet des consignes et des critères d'évaluation. À ce propos, nous pouvons nous déjà poser deux questions.

- Y a-t-il un lien entre la multiplication des types d'exercices et la multiplication des consignes ?
- Y a-t-il un lien entre la multiplication des types d'exercices et la multiplication des critères quantitatifs d'évaluation ?

En gardant ces deux questions à l'esprit, commençons par définir plus précisément la nature de cette évolution.

a) Les quatre types "traditionnels" d'énoncés dans les années 1980 et 1990

Dans les années 1980 et 1990, tous les énoncés auxquels nous avons eu accès (à quelques exceptions près) reflètent l'une de ces quatre formes fondamentales d'exercices : la dissertation ; l'analyse de texte ; des questions sur un texte ; des questions sur le cours.

Jusqu'au milieu des années 1990, la quasi-totalité des épreuves ne comporte qu'une seule des quatre formes d'exercices "classiques" ou, au plus, demande aux élèves de choisir l'une de ces formes. Dans le premier cas de figure, on demande à l'élève de répondre à deux ou trois questions sur le cours, d'analyser un texte ou de traiter un sujet de dissertation à choix parmi plusieurs sujets. Un exemple typique du deuxième cas de figure, qui constitue encore aujourd'hui le cadre de l'examen de philosophie du baccalauréat français³, est celui où l'on laisse l'élève choisir entre un sujet de dissertation (parmi deux sujets proposés) et un texte à analyser.

À partir de la deuxième moitié des années 1990, nous avons pu constater, à côté de ces quatre formes "traditionnelles", l'apparition de types d'énoncés sensiblement différents, témoignant de deux tendances majeures : d'un côté, une transformation de ces quatre types d'exercices (qui n'affecte pas de façon substantielle leur nature, mais plutôt la façon dont l'élève est amené, par l'énoncé lui-même, à effectuer la tâche en question) et, de l'autre, l'émergence de types d'exercices différents. Ces deux tendances, quoique distinctes, entretiennent des liens très étroits et participent d'un même mouvement général de multiplication des types d'énoncés et de fragmentation des contenus des épreuves.

b) Évolution des types "traditionnels" d'énoncés à partir de la deuxième moitié des années 1990.

À partir de la deuxième moitié des années 1990, on peut noter une transformation des types "traditionnels" dans trois directions : la cumulation de plusieurs types d'exercice, l'hybridation entre différents types d'exercice et la fragmentation d'un même type d'exercice.

--> La cumulation de plusieurs formes

On voit apparaître des épreuves comportant simultanément plusieurs types d'exercices. On trouve un exemple particulièrement significatif de ce procédé dans l'énoncé du Collège de Staël du 18 novembre 1998. L'épreuve est structurée en trois parties : connaissances et compréhension ; interprétation ; réflexion

On commence par neuf questions de compréhension d'un premier texte en relation avec le cours, pour passer à une brève explication d'un deuxième texte ("Dégagez le sens philosophique de cet extrait dans la perspective de la théorie de la connaissance de Platon - Environ une page -") et terminer, enfin, par une petite dissertation ("Proposez une esquisse - environ une page - de réflexion personnelle à partir de l'une des citations suivantes").

Une autre possibilité de cumulation nous est présentée dans un énoncé du Collège Calvin de juin 1999, comportant trois sujets : les deux premiers sont constitués de quatre questions portant sur le cours, alors que le troisième propose une question sur un texte. Une autre épreuve du même établissement, mais d'un maître différent, présente une question sur le cours et une explication de texte.

--> L'hybridation entre différentes formes

Dans un énoncé du collège Calvin de 2009, la forme "traditionnelle" de l'explication de texte n'est pas seulement cumulée avec d'autres formes d'exercices (ici, trois questions sur le cours), mais combinée avec elles. Trois textes assez courts (un fragment de Démocrite, un extrait d'Aristote sur Démocrite et un extrait de Diogène Laërce sur Démocrite) sont présentés aux élèves avec cette consigne : "Lisez attentivement les textes suivants ; dégagez-en brièvement les problématiques principales (une problématique par texte) ; expliquez brièvement quelle est l'opinion de Démocrite sur ces problématiques ; expliquez les textes (notions, arguments,...) et situez-les dans le contexte de la pensée de Démocrite ; comparez la pensée de Démocrite avec celle d'autres philosophes (École de Milet, Héraclite, Parménide) ; enfin, concluez par une réflexion personnelle sur les textes et leurs problématique. [Vous pouvez expliquer les textes séparément ou de façon groupée, comme vous le préférez]".

L'exercice, présenté comme une "explication de texte", est en effet dirigé de telle sorte que, en raison du caractère très détaillé des consignes, il est difficile de ne pas voir dans ces consignes mêmes des questions sur le texte. De plus, une partie de ces questions ("situez [les problématiques] dans le contexte de la pensée de Démocrite", "comparez la pensée de Démocrite avec celle d'autres philosophes") amène l'élève à mobiliser des ressources relatives aux contenus du cours et à exercer des compétences philosophiques bien précises ("dégagez [...] les problématiques", "comparez la pensée de Démocrite avec celle d'autres philosophes", "concluez par une réflexion personnelle").

Ce même énoncé présente donc également des traits typiques de fragmentation.

--> La fragmentation d'une même forme

Cette deuxième forme d'évolution concerne, en égale mesure, les quatre types d'énoncés "traditionnels" des années 1980-1990. En étudiant les énoncés constitués de questions portant sur les contenus du cours, nous avons observé que, dans les années 1980-1990, les questions posées sont rarement supérieures en nombre à deux ou trois, alors qu'à partir de la fin des années 1990 on trouve plus facilement des énoncés multipliant le nombre des questions. C'est par exemple le cas d'un énoncé du Collège Sismondi du 30 mai 2006 qui présente 12 questions.

On peut considérer ces formes d'énoncés comme des cas de fragmentation de la forme "traditionnelle", dans la mesure où l'on passe de quelques questions générales, demandant un long développement, à plusieurs questions plus ciblées, portant sur des fragments précis de la matière étudiée, et appelant à des réponses plus succinctes. Concernant ce dernier point, il arrive que le maître fixe, dans les consignes ou par la mise en page même de l'énoncé, le nombre maximum de mots pour chaque réponse, mais il arrive également qu'il signale leur nombre minimum.

Le même phénomène de fragmentation apparaît, à la même période, dans les énoncés comportant des questions sur un texte. Nous en avons un exemple dans une épreuve du Collège Voltaire de novembre 2001. Dans la première partie de l'épreuve, l'enseignant pose aux élèves sept questions sur un texte, questions qui permettent en quelque sorte d'en faire le tour. En d'autres termes, l'élève est guidé dans sa lecture du texte par l'enseignant, qui l'interroge en quelque sorte à sa place. Cela revient en dernier ressort à effectuer une analyse de texte dirigée.

Or, jusqu'au milieu des années 1990, l'explication de texte présente des consignes très succinctes, dont voici cinq exemples assez typiques : "Dégagez l'intérêt philosophique du texte suivant à partir de son étude ordonnée". "Faites une analyse critique du texte suivant en mettant en lumière l'enjeu de la problématique freudienne". "Voici un texte de René Schaerer. Analysez-le, puis prolongez si possible votre analyse dans une réflexion personnelle". "Voici un texte d'Henri Bergson. Faites-en l'analyse, en mettant en lumière les thèses essentielles qu'il contient". "Faites une analyse critique du texte ci-dessous".

Vers le milieu des années 1990, nous remarquons l'apparition de consignes plus précises et articulées, comportant un commentaire plus développé du maître, spécifiant davantage les étapes du travail à accomplir et allant parfois jusqu'à expliciter certains contenus (thème, problématique ou enjeux) du texte à analyser. On peut trouver un exemple très significatif de la première tendance dans cet énoncé de 1995 : "Lisez attentivement le texte ci-dessous, puis reportez-vous aux consignes de travail énoncées à la dernière page [Suit un extrait tiré de la République de Platon]. Vous procéderez maintenant de la façon suivante : Dans un premier temps, vous analyserez le dialogue entre Socrate et Glaucon en essayant de mettre en lumière les thèses essentielles qu'il contient. N'oubliez pas d'interpréter, chemin faisant, l'analogie métaphorique du rêve et de l'état de veille : que veut dire Socrate par-là ? - 2/3 pages - Socrate affirme, dans le Phédon, que le vrai philosophe ne saurait avoir peur de la mort, ni se révolter contre elle. Dites pourquoi - 1/2 pages-.

Dans un énoncé du même maître, datant de novembre 1999, le texte est précédé d'un titre explicitant le thème du passage ("Qu'est-ce donc que l'amour d'amitié (philia) ?") et suivi des consignes suivantes : "Après avoir soigneusement mis en lumière l'articulation du raisonnement de l'auteur, vous essayerez une comparaison entre la philia aristotélicienne et l'eros platonicien, puis vous vous efforcerez de développer votre propre point de vue".

Mais la part d'explication du texte, ou d'explicitation d'une partie de ses contenus, effectuée par le maître lui-même, peut être encore plus importante : "Kant attaque dans ce texte toute forme

d'eudémonisme. Dans un premier temps, vous expliquerez aussi clairement que possible quelles sont ses objections, notamment en quoi consiste le cercle sophistique et la contradiction qu'il dénonce. Vous direz ensuite comment un eudémoniste pourrait y répondre ; en guise de conclusion, vous exprimerez votre point de vue sur la question de la place du devoir et du bonheur en morale".

Nous avons pu observer une évolution similaire au sujet des énoncés de dissertation. Jusqu'au milieu des années 1990, les énoncés de dissertation comportent essentiellement deux éléments : des consignes, très succinctes, se limitant le plus souvent à rappeler le choix laissé aux élèves entre plusieurs sujets à "traiter", suivies de la liste, plus ou moins longue (de 3 à 5 sujets en moyenne), des sujets : "Dissertez librement sur l'un des sujets suivants" ; "Traitez, au choix, l'un des sujets suivants" ; "Traitez l'un des sujets ci-dessous" ; "Choisissez l'une des questions suivantes".

Les sujets proposés peuvent avoir deux formes différentes. Il peut s'agir d'une question directement posée à l'élève, par exemple : "Faut-il opposer l'art et la morale?". Il peut s'agir également d'une citation à discuter, par exemple : "Il n'y a qu'une ressource avec la mort : faire de l'Art avec elle" (René Char).

Or, partir de la fin des années 1990, nous observons l'apparition de consignes de plus en plus précises, visant à diriger davantage les élèves dans l'accomplissement de cet exercice et tendant à subdiviser, autant que possible, leur travail en tâches de plus en plus simples. Par exemple, un énoncé de janvier 2002 ajoute les précisions suivantes à la mention du sujet de dissertation ("La justice est la bonne organisation de la ville") :

"Présentez l'organisation de la ville proposée par Platon (dans La République) et la justification qu'il en donne. Êtes-vous d'accord ?"

De façon assez transparente, le but de ces précisions est, tout d'abord, d'amener les élèves à faire le lien entre le sujet et le cours, puis à structurer leur copie en deux parties : une première partie consistant en une restitution de ce même cours et une deuxième partie, critique, dans laquelle les élèves sont invités à discuter le sujet, en argumentant pour ou contre la théorie platonicienne.

Les consignes données par le maître peuvent être plus ou moins conséquentes et attirer l'attention de l'élève sur tel ou tel aspect de la dissertation. Par exemple, dans une épreuve de mars 1999, le maître effectue lui-même une partie du travail, en esquissant une analyse du sujet ou dégagant une série de demandes permettant de le questionner. Toutes ces stratégies, par lesquelles l'enseignant rappelle à l'élève les différentes opérations qu'il doit accomplir pour réussir l'exercice, voire en effectue certaines à sa place, opèrent une sorte de fragmentation de la forme même de l'exercice en ses composantes plus fondamentales.

Les énoncés se transforment donc dans le sens d'exercices de plus en plus dirigés et, par les consignes qui sont données ou la prise en charge, de la part du maître, d'une partie de l'exercice, déconstruits en leurs parties constitutives. Poussée jusqu'au bout, cette logique de fragmentation va

jusqu'à produire des nouvelles formes d'exercices, prenant en charge chacun une partie plus ou moins grande d'une forme "traditionnelle".

--> Apparition de nouvelles formes de question après 2000

Après 2000, nous pouvons noter, à côté des formes "traditionnelles" d'énoncés et de leurs formes dérivées, l'émergence de types d'épreuves sensiblement nouveaux. Nous avons essayé d'en esquisser une typologie, selon l'objet spécifique évalué par chaque type⁴ :

Évaluation des connaissances

Évaluation des compétences :

- argumentatives (argumenter)
- critiques (problématiser)
- conceptuelles (conceptualisation)
- évaluation des capacités d'analyse conceptuelle
- évaluation des capacités de synthèse conceptuelle

L'exemple le plus significatif de forme nouvelle du premier type est l'apparition, vers le milieu des années 2000, de QCM.

Dans les énoncés que nous avons pu consulter, le premier ayant cette forme remonte à 2004. L'épreuve est divisée en deux parties, dont la première ("Connaissance et compréhension") comporte trente questions à choix multiple sur les contenus du cours, tandis que la deuxième ("Interprétation et réflexion") demande à l'élève d'exposer, sous forme d'une lettre adressée à un ami, l'argumentation développée par Socrate dans un passage du Criton, et de développer sa propre position critique. Le QCM, visant à tester les connaissances relatives aux contenus du cours, est donc cumulé ici avec un exercice portant sur l'évaluation de compétences argumentatives.

On retrouve enfin des exemples de QCM dans une forme hybride, comme par exemple dans un énoncé du Collège Sismondi de 2005-2006. L'épreuve comporte six questions à choix multiple (mais cette fois il est demandé aux élèves de justifier "succinctement" leur réponse) et cinq questions "à développer". Parmi ces dernières, trois portent sur les contenus du cours, tandis qu'une autre consiste à expliquer une brève citation, et une dernière à analyser trois syllogismes : "Lesquels sont valides ? Lesquels sont composés de propositions vraies ?". Dans cet exemple, le QCM, visant essentiellement la validation des connaissances, intègre une part d'argumentation, et il est couplé avec des exercices ayant pour objectif l'évaluation de compétences lectorales et d'analyse logique.

À la même époque, on voit apparaître des épreuves ciblant exclusivement des compétences argumentatives et conceptuelles, comme dans un énoncé du Collège Rousseau de juin 2006. La

première partie est constituée de deux "Questions théoriques" : premièrement, on demande à l'élève de donner "le plan argumentatif détaillé (thèses et arguments)" d'un chapitre de l'Ethique d'Aristote, puis d'analyser méticuleusement chaque élément de la définition aristotélicienne de la vertu. La deuxième partie est constituée de deux "Exercices de logique", portant sur des syllogismes.

Mais il existe également des épreuves portant essentiellement sur l'analyse et la problématisation, comme cet énoncé du CEC De Staël de janvier 2003 : dans la première partie du travail ("Analyse et interprétation d'un document publicitaire"), l'élève doit analyser "le document (image et texte)", qu'il a préalablement choisi d'amener à l'examen, "pour en isoler les éléments significatifs", puis dégager "le sens" de chaque élément "pour [en] mettre en évidence les valeurs" relatives, et enfin élargir cette réflexion sur les "traits symboliques" du monde sportif aux "valeurs collectifs de la société actuelle". La deuxième partie de l'épreuve prolonge cet exercice de conceptualisation et de problématisation par une "esquisse de dissertation", en proposant quatre sujets au choix dont la mise en place présume les compétences exercées dans la première partie de l'épreuve.

Il est intéressant de remarquer l'apparition, ces dernières années, d'énoncés somme toute assez traditionnels quant à leur nature, mais innovants quant à leur forme, ou à la façon de présenter l'exercice "traditionnel" aux élèves.

Deux exemples : des images s'ajoutent parfois à des sujets "traditionnels" de dissertation voire, dans certains cas, les remplacent. Dans ces cas, on ne disserte plus (ou plus seulement) sur un sujet, mais sur une image. Ou bien, au lieu de demander aux élèves de rédiger une dissertation, de répondre à une question sur le cours ou sur un texte, on leur demande d'écrire un "essai philosophique", "une lettre à un ami", voire l'"éditorial d'un journal".

-->Tendance générale et nuances

Eu égard au caractère partiel et fragmentaire de notre corpus et à la grande variabilité des pratiques entre enseignants et établissements différents qu'on peut constater, il nous paraît difficile de pouvoir généraliser les résultats de notre analyse sur les types d'énoncés.

Ce que nous pouvons faire, par contre, c'est signaler l'émergence, surtout à partir de la deuxième moitié des années 1995, de nouvelles formes de l'évaluation écrite en philosophie. Ces formes attestent une tendance à une fragmentation et une multiplication des formes d'énoncés parallèles et, dans la plupart des cas, au morcellement des savoirs et des savoir-faire évalués.

Cela ne signifie pas que les formes plus "traditionnelles" d'épreuve aient disparu, car non seulement elles sont toujours présentes, mais elles restent probablement majoritaires. Le fait est, tout simplement, que des formes nouvelles apparaissent à côté des anciennes. Il s'agit maintenant de formuler quelques hypothèses susceptibles de rendre compte de cette émergence.

--> Hypothèses explicatives

De nouveau, on pourrait imputer cette tendance à une baisse des compétences correspondantes des élèves. Prenons l'exemple de la dissertation, qui a été, pendant plus d'un siècle, le paradigme même de l'épreuve scolaire de philosophie (Poucet, 2001). Cet exercice, demandant à l'élève de procéder à "l'étude méthodique et progressive des diverses dimensions d'une question"⁵, présuppose un degré de maîtrise élevé de plusieurs éléments⁶. On pourrait alors se demander si les élèves d'aujourd'hui maîtrisent encore tous ces éléments.

Publié en 1989, le Rapport de la Commission de Philosophie et d'Épistémologie, instituée en France par le Ministère de l'Éducation nationale et co-présidée par Jacques Bouveresse et Jacques Derrida, semble suggérer que cela ne soit plus le cas pour la plupart des élèves des classes terminales : "Dans les conditions actuelles, la plupart des copies de baccalauréat ne répondent pas aux exigences minimales d'une dissertation de philosophie et l'épreuve n'offre pas un instrument fiable d'évaluation des compétences effectivement acquises par les élèves. Pour de multiples raisons - la diversité illimitée des sujets, leur extrême généralité et leurs liens trop indirects avec ce qui a été étudié pendant l'année, l'appel à des capacités rhétoriques inaccessibles à la majorité des élèves actuels [...] - elle apparaît aux candidats comme mystérieuse et aléatoire ; son caractère non maîtrisable suscite l'angoisse, le bachotage, ou l'abandon et met peu à peu en cause l'enseignement philosophique lui-même" (p. 3).

Les épreuves "traditionnelles" seraient devenues trop difficiles pour des élèves "n'ayant plus le niveau" ou, pour le dire de façon plus neutre, n'ayant plus les compétences nécessaires pour réussir ces types d'exercices.

- Première hypothèse

Une première hypothèse explicative des phénomènes observés ferait donc le lien avec d'éventuelles causes exogènes et consisterait à dire que la différenciation des pratiques évaluatives serait une réponse des enseignants à une baisse de compétence et de motivation des élèves. Une première possibilité consisterait à réorganiser l'évaluation en fonction de tâches plus simples et mieux maîtrisables.

Dans ce cas, les enseignants auront la tendance soit à formuler des énoncés tendant à mieux diriger les élèves à travers une tâche complexe, soit à simplifier une tâche complexe, en effectuant eux-mêmes une partie du travail, soit à élémentariser les tâches complexes en tâches plus simples. Quoi qu'il en soit, ces différentes stratégies semblent refléter une même volonté de favoriser le sentiment de maîtrise des élèves, ce qui permettrait également de diminuer leur angoisse et de mieux gérer leur peur d'échouer.

Reprenons l'exemple de la dissertation. À plusieurs égards, on pourrait considérer la dissertation comme étant l'exercice de la maîtrise par excellence. Or, dans une certaine conception (magistrale) de l'enseignement, cette maîtrise est justement l'affaire du maître qui, concevant sa leçon comme une longue dissertation, se présente face à ses élèves comme un modèle à imiter (Muglioni, 2002).

L'apparition de nouvelles formes d'énoncés, ciblant des compétences spécifiques, pourrait s'expliquer par la prise de conscience de la part de certains enseignants des limites d'un modèle strictement transmissif, et de son abandon en faveur de choix didactiques différents, tels qu'il peuvent être suggérés par les modèles béhavioristes, cognitivistes ou constructivistes, portant une attention particulière au rôle des renforcements positifs dans les dynamiques d'apprentissage, à leur régulation, à l'adaptation des objectifs pédagogiques aux compétences réelles des élèves, à la différenciation des activités didactiques etc.

L'application de modèles d'apprentissage différents impliquerait alors l'adoption de paradigmes d'évaluation différents. Pour reprendre les termes de Ketele (1993), le paradigme de l'intuition pragmatique, considérant l'évaluation comme un "acte intuitif", laisserait la place à d'autres modèles, tels que l'évaluation centrée sur les objectifs, ou le modèle bloomien de la maîtrise, beaucoup plus attentifs à la validation progressive de savoirs et de savoir-faire faisant l'objet des différentes séquences d'enseignement, et à la dimension formative de l'évaluation. L'apparition de QCM, ou la multiplication des questions dans un énoncé, témoignerait également du souci docimologique pour une plus grande objectivité de l'évaluation. Ces pratiques favoriseraient en outre le respect de la règle des 2/3, prescrivant de "donner à l'élève au moins trois occasions indépendantes l'une des autres de démontrer sa compétence" (Gerard, 2008, p. 80). Ces stratégies contribueraient ainsi à combattre une désaffection possible des élèves pour la philosophie, s'il est vrai, comme le montrent Tardif (1992) et Vianin (2006), que la "perception de la contrôlabilité de la tâche" a une influence importante sur motivation.

Une deuxième possibilité, visant directement la motivation des élèves, consisterait à travailler directement sur la "perception de la valeur de la tâche" (ibid.). L'apparition d'énoncés comportant des images, ou proposant des formes d'écriture moins "traditionnelles" (par exemple la lettre ou l'éditorial), ou portant sur des thèmes d'actualité (la publicité, le sport etc.), refléterait la tentative de certains enseignants de renouveler la forme "traditionnelle" d'épreuve, par exemple par leur ancrage dans des "questions socialement vives" (Chevallard, 1997). Ce procédé semble renvoyer de façon assez transparente à une volonté plus générale de redonner un sens, voir "une saveur", au savoir enseigné (Astolfi, 2008).

- Deuxième hypothèse

Une deuxième hypothèse endogène permettant d'expliquer une partie des phénomènes observés ferait référence non plus à l'angoisse des élèves confrontés à des exercices perçus comme "non maîtrisables", mais à celle d'enseignants n'arrivant plus à maîtriser l'apprentissage de leurs élèves. Il se pourrait en effet que l'enseignant, dans certains cas, au lieu d'adapter ses stratégies

d'enseignement aux élèves afin de mieux réguler le processus d'apprentissage, rendent simplement ses épreuves "plus faciles", étant dans l'incapacité de trouver des stratégies efficaces pour remédier à ce qu'il perçoit comme une "baisse de niveau" générale des apprenants. Il se pourrait également que l'enseignant, tout en proposant aux élèves une épreuve "traditionnelle", garde le même niveau d'exigence au niveau de la tâche demandée, mais adopte la même stratégie d'adaptation à la baisse, en étant moins exigeant au moment de la notation.

- Troisième hypothèse

La transformation des types d'énoncés et l'apparition de nouvelles formes d'exercices pourrait encore être comprise comme un élément participant à un processus plus général de transformation du travail de l'enseignant au sein de l'école. Nous ne pourrions pas, en d'autres termes, séparer la question de la nature des énoncés de celle des consignes et des critères d'évaluation : il s'agirait de phénomènes solidaires, reflétant un même processus et se renforçant les uns les autres.

Il est indéniable que l'introduction de nouveaux types d'exercices, visant des compétences de plus en plus ciblées et spécifiques, favorise la multiplication des consignes. Mais un rapport de cause à effet inversé paraît, comme nous avons essayé de le montrer, également vraisemblable : la multiplication des consignes semble avoir eu pour effet le morcellement de certaines formes d'exercices, jusqu'à l'émergence de formes décidément nouvelles.

Le même raisonnement peut être fait au sujet de la multiplication des critères d'évaluation. D'un côté, la fragmentation et la multiplication des exercices à l'intérieur d'une même épreuve favorisent une meilleure prise en charge "comptable" de l'évaluation (attribution de points pour chaque question, élaboration de barèmes etc.). D'un autre côté, la pression exercée sur les enseignants pour rendre plus transparente et "objective" l'évaluation a peut-être pas favorisé, à son tour, une fragmentation des épreuves.

Les mêmes hypothèses formulées plus haut, dans le but de rendre compte de la multiplication des consignes et de l'apparition massive de critères quantitatifs d'évaluation, pourraient donc contribuer à l'explication de l'évolution des types mêmes d'énoncés : la réaction des enseignants à une perte objective de certaines compétences des élèves (1), que nous avons déjà évoquée, le besoin de mieux contrôler les élèves (2), celui de prévenir leur échec (3) et de fixer des critères pour prévenir leurs contestations ou celles de leurs familles (4).

- Quatrième hypothèse

Nous pourrions aussi formuler une nouvelle hypothèse, faisant écho à l'hypothèse (2) : le besoin de l'enseignant de mieux contrôler, non pas les élèves, mais ses propres pratiques d'enseignement et d'évaluation. En d'autres termes, la production d'énoncés davantage ciblés sur des objectifs facilement identifiables, permettrait à l'enseignant non seulement de mieux réguler l'apprentissage des élèves, mais également de vérifier l'efficacité de son dispositif didactique, favorisant ainsi un

retour réflexif sur son propre travail et les objectifs qu'il s'est fixés à lui-même. Dans ce cas, on pourrait dire que l'enseignant choisit parfois d'évaluer différemment pour mieux s'auto-évaluer. Une partie de l'évolution de pratiques d'évaluation des enseignants de philosophie du secondaire genevois pourrait, en ce sens, être lue comme solidaire au processus de professionnalisation de l'enseignement, favorisant le modèle du "praticien réflexif" (Maroy & Cattonar, 2002).

- Cinquième hypothèse

Enfin, l'apparition de nouvelles formes d'énoncés pourrait en partie s'expliquer par l'apparition de nouveaux objets d'enseignement. En particulier, l'une des raisons susceptibles d'expliquer l'émergence de certains énoncés ciblant de façon quasi exclusive les compétences argumentatives et de raisonnement logique, semblerait d'ordre essentiellement épistémologique : il s'agirait de l'impact, sur les pratiques enseignantes de certains maîtres de philosophie, d'un courant philosophique de plus en plus présent au niveau académique (la philosophie dite "analytique").

Cette hypothèse nous paraît intéressante dans la mesure où elle nous amène à poser la question de l'influence des options idéologiques d'un enseignant (pour nous, sa préférence pour un courant philosophique spécifique, mais aussi pour un auteur, une époque, une thématique etc.) sur ses choix didactiques et, plus largement, sur les objectifs pédagogiques qui orientent son action. Le choix de tel ou tel autre type d'épreuve par un enseignant semble en effet en partie également tributaire du savoir qu'il se propose d'enseigner, ainsi que des représentations et des valeurs qui lui sont plus ou moins sous-jacents.

III/ Troisième Partie : résultats et conclusions

L'analyse du matériel que nous avons recueilli nous a permis de mettre en évidence les transformations subies, au cours des trente dernières années, par les énoncés d'évaluation dans les épreuves écrites de philosophie. Nous nous sommes concentrés sur trois d'entre elles, qui nous paraissent particulièrement significatives : la multiplication et la diversification des consignes ; l'apparition et la multiplication des critères d'évaluation, et notamment des critères quantitatifs ; la fragmentation et la différenciation des types d'énoncés. L'étude de ces trois points nous a enfin conduits à la formulation d'hypothèses susceptibles de rendre compte des phénomènes observés, travail qui, à son tour, nous a permis de nous interroger sur plusieurs dimensions de la profession enseignante, et sur les causes possibles de son évolution au cours des trente dernières années. Pour conclure, nous souhaiterions revenir sur quelques-uns de ces aspects, pouvant nourrir avec profit une réflexion sur nos propres pratiques enseignantes, ainsi que sur le contexte scolaire et social dans lequel elles s'intègrent.

- La première remarque concerne la tendance, que nous avons pu remarquer, à une plus grande transparence des consignes et des critères d'évaluation. Nous pouvons nous poser la question : est-ce le fait des élèves, ou bien des maîtres, ou bien d'une sensible transformation de la place même de l'école au sein de la société ? Qu'il y ait eu, au cours de ces trente dernières années, une pression

croissante de la part des élèves et de leurs familles sur les enseignants et l'école en général, en termes de résultats scolaires, cela semble indéniable. La volonté, de la part des familles et de l'institution scolaire d'exercer un contrôle accru sur le travail des enseignants, pourrait en ce sens témoigner d'une fracture qui s'est produite au niveau même de la confiance entre les différents acteurs scolaires. L'école et la société peuvent-elles encore se permettre de faire confiance à un enseignant qui fonde son évaluation sur sa seule "intuition" et son "expérience" ? N'est-il pas préférable que l'évaluation s'appuie sur des critères totalement objectifs ? Mais cette objectivité absolue n'est-elle pas à son tour, et pour ainsi dire, un mirage docimologique ? Quelle place peut-on encore attribuer, dans l'école d'aujourd'hui, aux "savoirs d'expérience" des enseignants (Tardif, Lessard & Lahaye, 1991) dans le domaine de l'évaluation ?

Bien sûr, nous avons déjà eu l'occasion de signaler la possibilité que l'enseignant, dans un souci d'efficacité didactique, intériorise cette exigence de contrôle. Mais nous pouvons légitimement nous demander si une demande excessive de contrôle exercé sur le travail des enseignants n'aurait pas pour conséquence de rendre encore plus fragile la confiance des élèves envers les enseignants, voire de mettre à mal l'autorité de ces derniers, ce qui nuirait à l'efficacité de leur action pédagogique.

- La deuxième remarque concerne la baisse de compétences des élèves, qui permettrait d'expliquer une bonne partie des transformations ayant trait aux consignes et à la nature des énoncés. Au fil de notre travail, cette hypothèse nous a conduit à nous interroger sur les stratégies adoptées par les enseignants pour faire face à cette nouvelle réalité, stratégies qui permettraient d'expliquer une partie de l'évolution de leurs pratiques d'évaluation. Or, que montre-t-elle, en creux, cette évolution ? Est-ce le besoin croissant, manifesté par les élèves, à se faire guider qui pousse le maître à multiplier des consignes de plus en plus précises ? Ou alors faut-il expliquer ce phénomène par le besoin de l'enseignant lui-même de prévenir l'échec de ses élèves, afin de se rassurer sur la valeur de son enseignement ? Dans le cadre de ce travail, il nous est difficile d'exclure que cette évolution ne puisse cacher une baisse du niveau d'exigence de certains enseignants, démunis face à une baisse de niveau des compétences de leurs élèves. Mais il nous semble possible d'affirmer qu'elle nous montre clairement une hausse du niveau des pratiques des enseignants, reflétant leur travail pour faire face aux difficultés manifestés par les élèves et peut-être, aussi, à leur propre "lassitude devant la charge de l'évaluation" (Troger, 2005).

Il serait dès lors impossible de comprendre le travail accompli dans le domaine de l'évaluation indépendamment de la considération du rôle qui lui est attribué au sein du dispositif pédagogique général mis en place par l'enseignant. C'est du moins ce qui nous a paru évident lors de l'analyse d'énoncés qui semblent être l'expression d'une action pédagogique axée sur des objectifs précis, visant non seulement à rendre plus fiable l'évaluation, mais également à améliorer la régulation des apprentissages et à favoriser un retour critique de l'enseignant lui-même sur l'efficacité du dispositif adopté et de ses propres pratiques pédagogiques. La tendance, que nous avons pu remarquer, à l'hétérogénéité des pratiques d'évaluation, serait donc le reflet d'une différenciation pédagogique

mise en oeuvre par les enseignants dans le but de faire face à l'hétérogénéité des publics scolaires (Meirieu, 1988).

En conclusion, les résultats de ce travail nous invitent à poursuivre notre questionnement au-delà du cadre que nous lui avons donné en formulant le carré docimologique, voire au-delà du cadre fixé par le carré lui-même. Les trois aspects que nous venons d'évoquer nous poussent en d'autres termes à nous interroger sur le contexte social, institutionnel et épistémologique dans lequel les différentes pratiques d'évaluation sont mises en oeuvre et qu'on pourrait figurer comme le cercle dans lequel s'inscrit le carré.

(1) En 1930 le professeur Laugier effectue une expérience de multicorrection de copies d'agrégation d'histoire puisées dans les archives. 166 copies sont corrigées par 2 professeurs travaillant séparément, sans connaître leurs appréciations respectives. Tous deux ont une longue expérience et corrigent méticuleusement. Les résultats sont surprenants. La moyenne des notes du premier correcteur dépasse de près de deux points celle du second. Le candidat classé avant dernier par l'un est classé second par l'autre. Les écarts de notes vont jusqu'à 9 points. Le premier correcteur donne un 5 à 21 copies cotées entre 2 et 14 par le second, le second donne un 7 à 20 copies cotées entre 2 et 11,5 par le premier. La moitié des candidats reçus par un correcteur est refusée par l'autre. En 1932, la Commission Carnégie effectue une expérience de multicorrection en prélevant, au hasard, cent copies dans les archives du baccalauréat à Paris. Ces copies sont distribuées à 6 groupes de 5 examinateurs. Les disciplines concernées sont le français, la philosophie, le latin, les mathématiques et la physique. On demande aux examinateurs de noter les copies et de fournir un rapport sur (1) les qualités exigées, (2) le classement des dites copies - la méthode de notation utilisée. Les résultats montrent une forte dispersion des notes attribuées à chaque copie par les correcteurs. Aucune copie ne reçoit deux fois la même note. L'écart maximum des notes dépasse les prévisions. Une copie de français est notée 3 et 16. En philosophie et en latin, l'écart maximum est de 12 points. Les mathématiques et la physique, réputées pour des sciences exactes, ne sont pas épargnées : l'écart maximum est respectivement de 9 et 8 points. D'autre part, le classement des copies variait fortement d'un correcteur à l'autre (Naumier, 2006).

(2) Précisons que nous ne l'avons pas conçue à partir de modèles théoriques, mais qu'elle nous a été suggérée par un professeur de français et philosophie proche de la retraite. On lui a demandé de faire une sorte de bilan de sa longue carrière d'évaluateur, et il a communiqué ceci (Collège de Staël, communication privée du 18 mars 2010) : "Autrefois (...) les professeurs de philosophie proposaient volontiers des sujets de dissertation ; peu à peu ils se sont mis à privilégier les explications de textes sans "guidage". Puis ces textes à expliquer ont été assortis d'un questionnaire, mais les questions ne comportaient pas de points pour l'évaluation. En France, les épreuves du bac se sont peu à peu transformées en ce qu'elles ont proposé des questions à points. Certains profs ici ont suivi cette pratique. Il me semble que progressivement on en est venu à un type de plus en plus "comptable" de l'évaluation. L'évaluation intuitive du prof, portant sur l'ensemble du travail paraît aujourd'hui

obsolète. Il faut, aujourd'hui, sortir sa calculette pour additionner ou soustraire les points attribués ou retirés ! Analyse de ce phénomène sur la grille Boltanski (...) : on est passé des valeurs du monde 'inspiré', intuitif, etc. (...) au monde 'industriel' (rationalisé par l'exactitude comptable). Pourquoi cette évolution ? Peut-être par conformité au monde technologico-rationnel : tout doit pouvoir se traduire en chiffres (ce dont le monde "inspiré" se moque éperdument !). On donne ainsi l'impression d'éviter l'arbitraire. On se rapproche de l'idéal d'objectivité des QCM. Souci de maîtrise ! Plus prosaïquement : pour éviter les recours (à la maturité par exemple), par ailleurs de plus en plus nombreux dans notre République !".

(3) Note de service n°2001-154 du 30 juillet 2001 ("Définition de l'épreuve écrite et de l'oral de contrôle du baccalauréat général"), BO n°31 du 30 août 2001, Education nationale (France).

(4) Nous reprenons ici la classification des processus de pensée philosophique formulée par Tozzi (2005).

(5) "de philosophie en classe terminale des séries générales", BO n°25 du 19 juin 2003, Education nationale (France).

(6) Entre autres, à titre indicatif, on pourrait évoquer : la méthodologie propre à cette forme d'exercice ; les compétences spécifiques permettant de développer une pensée philosophique (par ex. conceptualiser, argumenter etc.) ; des compétences transversales, qui ne font pas directement l'objet des cours de philosophie, mais sont présupposées comme ayant déjà été acquises par les élèves pendant leur scolarisation antérieure (notamment, les compétences rédactionnelles) ; une connaissance suffisamment riche de la culture philosophique (et générale) ; une aptitude critique et réflexive permettant de problématiser les savoirs acquis et de se questionner sur ses propres démarches de pensée.