

Comment évaluer une pensée en philosophie ?

Guillaume Lequien, professeur de philosophie

I) Les contraintes et les préjugés qui pèsent sur l'évaluation en philosophie

En France, l'enseignement de la philosophie offert à tous les élèves des séries générales et technologiques est le croisement de plusieurs difficultés que la plupart des autres disciplines ne connaissent pas :

- en l'absence d'années préparatoires et d'années de consolidation en amont et en aval de la Terminale, tout le programme de philosophie fait l'objet d'une familiarisation et d'un apprentissage restreint sur une unique année scolaire.
- la rencontre avec la philosophie se confond très rapidement avec la préparation à l'épreuve de philosophie du baccalauréat, sans laisser de possibilité pour d'autres types d'écrits moins complexes, et avec un enjeu social et symbolique fort.
- sur cette seule année, les élèves entretiendront un rapport affectif avec la matière qui résultera pour moitié de leur propre intérêt intellectuel, et pour moitié des quelques notes qu'ils obtiendront au fil de l'année, couronnées par la note de philosophie au bac qui marque et sanctionne définitivement le regard qu'ils conserveront sur cette matière plus tard.

Un enjeu pédagogique fort est donc celui de l'évaluation en philosophie, sur lequel circulent plusieurs préjugés persistants :

- préjugé du recrachage de cours. Comme dans les autres matières, pour avoir une bonne note, il faut reproduire le cours en citant in extenso les grands auteurs qui ont été vus en cours. C'est ce préjugé qui normalise peu à peu la philosophie par rapport aux autres disciplines, et qui a pour revers l'abus des corrigés prêts à l'emploi dans les copies ;
- préjugé de l'originalité comme marqueur de la pensée. Pour avoir une bonne note, il faut présenter la copie la plus originale qui soit, quitte à ne pas produire la dissertation attendue : la légende urbaine du "l'audace, c'est ça" [= de rendre audacieusement une copie blanche au lieu de réfléchir sur ce qu'est l'audace] ;
- préjugé des affinités électives du correcteur. Pour avoir une bonne note, il faut que les idées présentes dans la copie plaisent au correcteur, donc qu'elles rejoignent ses propres opinions personnelles sur la liberté, la politique, le bonheur etc. ;
- préjugé de l'arbitraire du correcteur, et du désaccord profond entre les correcteurs. De toute façon, la note apposée sur la copie varie fortement et dépend en dernier recours de la personne du correcteur, et un autre correcteur aurait sensiblement mis une note différente.

II) Pourquoi évaluer une démarche de pensée ?

La présence de l'examen à la fin de la classe de philosophie impose une évaluation, sur une période d'apprentissage courte (neuf mois) et dense (car bien d'autres matières seront aussi évaluées). Cette exigence institutionnelle crée l'illusion qu'il est "naturel" pour un enseignant d'évaluer une production écrite d'un élève.

Or cette production écrite est une tâche complexe imposée aux élèves, souvent novices : il s'agit pour eux de montrer qu'ils comprennent une question qu'ils n'ont pourtant pas formulée eux-mêmes, et de lui apporter ensuite non pas une réponse argumentée, mais plusieurs réponses argumentées qui se positionnent sur un mode "dialectique" les unes par rapport aux autres, en mimant un effort de pensée qui, non content d'aboutir à une réponse, est capable de revenir sur cette réponse pour en critiquer les présupposés et nourrir ainsi une autre réponse, elle aussi argumentée. Cette exigence formelle, lourde, s'applique à des questions pouvant porter sur quasiment toute réalité possible, en théorie restreinte par les dimensions gigantesques d'un programme de notions qui ouvre la porte à une quasi infinité de questions. Quatre heures suffiront pour mesurer la réussite des meilleurs face à ce défi.

Outre ces difficultés propres à la situation de la philosophie dans l'enseignement secondaire, il y a une première question qui ne va pas de soi : qu'évalue-t-on exactement en philosophie, et pourquoi l'évalue-t-on ?

Évalue-t-on des idées, même argumentées ? Évalue-t-on la forme argumentative, la pertinence, la "qualité" de suggestion d'un discours ? Cela aurait-il un sens de mettre un 11 à Nietzsche car ses aphorismes sont suggestifs mais manquent de développement, un 7 à Socrate parce qu'il prend trop d'exemples personnels, un 18 à Sartre ou Heidegger parce que c'est dialectiquement pertinent, mais on pourrait leur reprocher néanmoins d'abuser de certains néologismes obscurs et pompeux ?

Évalue-t-on les différentes réponses apportées à la question ? Cela supposerait que certaines réponses sont attendues, soit par le correcteur lui-même, soit par le corps des enseignants.

En supposant qu'il faille évaluer le cheminement d'une pensée argumentée en réponse à une question préalable et mobilisant une certaine culture philosophique initiale, pourquoi l'évaluer ?

L'acte d'évaluation est banal pour l'enseignant qui doit fournir pour chaque élève de chaque classe un minimum de notes par trimestre, mais pour l'élève singulier qui a tenté de participer à ce nouveau jeu, et qui parfois a passé beaucoup trop d'heures pour accoucher d'une production écrite dont il ne comprend pas encore les défauts car il n'en maîtrise pas encore les attendus, la première évaluation reçue a un pouvoir symbolique et affectif fort, qui pourra déterminer en partie la relation ultérieure avec la philosophie. Avoir un 8/20 ou avoir un 14/20 au premier devoir de philosophie, quel qu'en soit le contenu exact que l'élève ne relira probablement pas, cela sanctionne d'un geste toute la qualité de sa pensée, réduite à un chiffre plus ou moins satisfaisant.

Or cet acte d'évaluation de la pensée d'une personne par une autre personne est une forme de violence qui ne va pas de soi, mais légitimée par la subordination de l'élève (ignorant) à l'enseignant (savant). En-dehors de l'institution scolaire, personne n'aurait l'idée de chiffrer quantitativement la "qualité" d'une pensée sur un horizon obscur menant à 20. Et pourtant cette évaluation est nécessaire, pour que l'élève prenne conscience de ses lacunes et progresse. Qu'une pensée doive se confronter à un regard extérieur pour éprouver sa propre pertinence, c'est un décentrement nécessaire, mais violent, surtout quand il s'agit d'entrer dans une nouvelle discipline.

L'un des dangers de l'évaluation consiste à évaluer chaque écrit pour finalement normaliser la pensée de l'élève, selon des normes qui sont en partie exigées par l'institution (les instructions officielles), et en partie interprétées et réordonnées par chaque enseignant, dans sa pratique propre. L'enseignant doit le premier montrer sa capacité à pouvoir se décentrer de son propre arbitraire, en n'attendant pas un devoir-type, et à accueillir avec bienveillance toute méthode, toute référence, tout cheminement qui ne seraient pas les siens. Mais cet arbitraire de l'enseignant, décrié par des générations d'élèves, n'est pas qu'un mythe : il est constitutif de l'acte même d'évaluer.

L'évaluation est ainsi biaisée deux fois : d'abord dans l'activité de correction elle-même où différents effets peuvent fausser l'appréciation de chaque copie singulière (c'est le "triangle pervers" de la correction, qui combine des effets liés au paquet de copies, des effets liés au correcteur, et des effets liés à l'élève ou à la perception qu'on en a) ; puis dans les normes explicites et implicites auxquelles j'essaie de me rattacher au fil de la correction.

III) Quand j'évalue une copie, quelles sont mes normes ?

J'évalue en fonction de mon propre parcours, des méthodes scolaires que j'ai acquises, des évaluations que j'ai reçues en tant qu'élève/étudiant, de la spécialisation qui a été la mienne, et qui m'a rendu plus sensible, soit à la rigueur analytique d'une pensée, soit à l'interprétation d'aphorismes profonds, etc.

J'évalue en fonction du professeur que je suis actuellement, je suis donc sensible au fait de retrouver dans une copie une référence que j'utilise en cours, je comprends moins une méthode d'un autre collègue qui transparait dans la copie et j'ai plus de mal à l'évaluer, par manque d'habitude.

J'évalue en fonction d'un certain idéal-type de la copie que j'attends, celle à laquelle je mettrai 20 sans hésitation ni remords, qui correspond peut-être à celle que je pense pouvoir produire si j'étais élève, ou même que j'aurais effectivement produite à l'époque où j'étais élève. Cet idéal-type n'a par définition aucune référence extérieure attestée, il résulte du savant mélange entre les maîtres qui m'ont impressionné et l'élève idéalisé que je pense avoir pu être.

J'évalue en fonction des autres élèves, des autres copies, insensiblement chaque copie est jugée relativement aux autres, et toutes sont mis en concurrence. Mais pour pouvoir comparer les devoirs, il faut nécessairement objectiver certains critères normatifs.

Je dois aussi évaluer en fonction des attentes supposées de l'institution, de ce qu'un correcteur idéal (que je ne suis pas) est prêt à admettre ou non dans une copie. La notation croisée des copies des autres m'amène aussi au bac à me décentrer légèrement de ma notation habituelle, dans un sens ou dans l'autre.

Mais :

1. Puis-je vraiment faire abstraction du professeur que je suis devenu quand j'évalue une copie d'examen ?
2. Quelles sont exactement les attentes du correcteur idéal ? Il n'est proposé aucune référence extérieure attestée, puisque l'institution elle-même tente d'harmoniser les notations sans prescrire de consignes de correction explicites.

Certaines expériences menées en formation ont ainsi permis de comparer la correction croisée de copies de Terminale par des enseignants de Terminale et par des enseignants de Licence de Philosophie : les enseignants de Terminale, dans un cadre généraliste, visent un niveau bien supérieur à l'enseignement supérieur spécialisé lui-même (une copie qui atteint difficilement la moyenne au bac la dépasse largement en première année de licence !). Le modèle formel de la dissertation vient des concours de recrutement d'enseignants (Est-ce comparable à un examen généraliste sanctionnant le niveau d'élèves en fin de secondaire ?), et le mode d'évaluation n'en est pas si éloigné qu'il devrait l'être.

Parmi toutes ces normes qui rendent possible l'évaluation, certaines sont conscientes, et d'autres non. Plus grave, à mon sens : certaines sont visibles et explicitement exigées, d'autres restent invisibles, dans les coulisses du travail de correction, et neutralisent en conséquence la prise de conscience par l'élève de ce qui ne va pas. Comment lever les implicites ? Comment l'évaluation peut-elle passer d'une arme de culpabilisation à une méthode de conscientisation ?

IV) Comment l'élève peut-il comprendre l'évaluation de sa copie ?

L'évaluation d'une copie pour laquelle ne sont fournis ni barème, ni critères d'évaluation explicites est injuste et inutile. Injuste car elle maintient l'élève dans l'obscurité sur ce qui lui fait encore défaut ; inutile car l'élève reproduira sensiblement les mêmes défauts, à moins de renoncer directement.

D'où la double nécessité pour l'enseignant-correcteur : objectiver ses normes de correction pour lui-même (pour échapper à ses propres biais inconscients), et pour les élèves (pour échapper au préjugé de l'arbitraire du professeur et rendre possible un apprentissage progressif).

V) Autre difficulté : peut-on concevoir et mettre en place une évaluation progressive sur neuf mois ?

La dissertation est une tâche complexe qui suppose la mobilisation de plusieurs savoirs et savoir-faire, dont l'acquisition demande un minimum de temps. Peut-on imposer dès septembre la rédaction d'une dissertation intégrale dont les codes n'ont pas encore été maîtrisés ?

Présenter une méthode puis en demander l'application immédiate, sans droit à l'erreur, est inhumain. Certes, les premières copies ont le droit de ne pas être parfaites du premier coup, mais cela signifie-t-il que le 20/20 est implicitement inaccessible en début d'année, et que par conséquent la notation sur 20 points est mensongère ? Si on accorde un droit à réaliser une copie imparfaite, est-il normal d'utiliser un système d'évaluation comparable à celui de l'examen final ?

Peut-on diviser la tâche complexe en plusieurs micro-tâches discrètes, et ainsi sur plusieurs mois amener jusqu'à l'achèvement d'une première dissertation ? Mais alors, comment évaluer ces micro-tâches ? Comment mesurer la qualité d'une dissertation entière que je ne lis pas en me basant uniquement sur une introduction que je lis ? Si une progressivité des exercices est mise en place, elle implique une progressivité de l'évaluation. Mais comment assurer la cohérence de l'évaluation de la partie au tout, si les deux mettent en jeu à peu près le même arbitraire sans explicitation ?

L'évaluation est chiffrée à l'examen, cela signifie-t-il qu'elle doit l'être constamment ? Evaluer, est-ce nécessairement noter ?

La note est l'aboutissement, quantitatif, d'une évaluation elle-même complexe, qui commence par des annotations marginales, et dont les points essentiels transparissent dans l'appréciation finale. L'appréciation précède et justifie la note. Mais du côté de l'élève, la copie est reçue dans l'ordre inverse : la note chiffrée, directement lisible, sanctionne la copie, et l'appréciation ne sera lue qu'ensuite, si elle est lue. La note invisibilise finalement l'appréciation dont elle est pourtant le prolongement.

Est-il concevable d'évaluer la copie sans la noter ? La note répond aux attentes des élèves, de leurs parents et de l'institution scolaire, alors que c'est l'évaluation qui devrait prévaloir, pour favoriser une progressivité de l'apprentissage.

Si l'on accorde un droit à l'erreur, ne pas se contenter de faire la moyenne des différents travaux quelle que soit leur situation dans la période donnée. La moyenne aplatit et rend invisibles les progrès sur des points précis.

VI) Evaluer en étant transparent à soi-même et à ceux que l'on évalue ? Une proposition soumise à la discussion

Quand j'ai commencé à enseigner, je me suis heurté à cette difficulté d'une évaluation impossible mais nécessaire, et j'ai trop souvent eu l'impression de subjectivité extrême impossible à partager et difficile à justifier quand je vois la façon dont certains collègues attribuent une note chiffrée "au feeling", feeling qui s'est constitué avec la pratique certes, mais qu'il serait injuste d'imposer aux élèves sans aucune explicitation. D'où mon objectif depuis : tenter de clarifier pour soi-même ses propres critères pour ne pas trop hésiter en cours de route, et les expliciter dès le départ pour les élèves.

Selon moi, pour évaluer dans la transparence, il faut commencer par identifier des compétences-clés dans un travail de philosophie :

- des compétences que l'on peut retrouver dans tous les types d'exercice, pour pouvoir les comparer ;
- des compétences dont on peut rendre visible l'acquisition sur un temps de quelques mois ;
- des compétences sur lesquelles les autres enseignants-correcteurs pourraient se mettre d'accord.

Où chercher ces compétences ?

Voici ce que prescrit le programme officiel : "apprentissage de la réflexion philosophique, aptitude à l'analyse, aptitude à utiliser les concepts élaborés et les réflexions développées ainsi qu'à transposer dans un travail philosophique personnel et vivant les connaissances acquises par l'étude des notions et des oeuvres". Est-il possible de transposer ces expressions sous la forme de compétences clairement définies ?

On peut également faire appel aux compétences définies par Michel Tozzi (2005) :

- savoir conceptualiser
- savoir problématiser
- savoir argumenter

Ou encore à d'autres séries de compétences, l'important étant de clairement les expliciter dès le départ et de constamment les évaluer de la même façon.

VII) Comment évaluer ces compétences ?

Éviter le barème chiffré.

Rendre visible l'évaluation : remplacer les chiffres par un code, de smileys s'il le faut, pour ma part j'ai opté pour un code de plus / moins qui vaut ce qu'il vaut.

Assurer la transposition de cette évaluation dans une note chiffrée : par exemple (cela fonctionne bien en séries technologiques) j'applique une règle de calcul simple, je hiérarchise les copies en fonction du nombre de + (sans compter les -) pour valoriser leurs points forts, puis j'additionne les + et les - et je corrèle avec une grille chiffrée de 0 à 20 où un total de 0 (exemple un + et un -) équivaut à une note chiffrée de 10, et j'échelonne en-dessous et au-dessus en fonction du total relatif et du nombre de +, ça permet à la fois d'être équitable sur un grand nombre de copies et d'en valoriser certaines sans que ce soit absolument arbitraire. En comparant leurs copies, les élèves voient que ça colle assez bien. La règle de calcul peut paraître fastidieuse en elle-même et discutable sur le principe, le tout est encore une fois de s'y attaquer en la pratiquant et en trouvant des variantes.

Cette grille de critères ne remplace pas l'annotation manuscrite globale, mais elle a le mérite de la faciliter (je remplis d'abord les + et les - puis j'explique verbalement) et de la raccourcir (inutile de répéter tout ce qui va / ne va pas dans le devoir, ça apparaît déjà dans la grille, je ne retiens que l'essentiel, rarement plus de trois lignes, c'est-à-dire rarement plus que ce que l'attention critique d'un élève sur son propre travail lui permet de retenir à la fin). La grille et l'annotation ne font pas double emploi et n'alourdissent pas la correction, cela prend 20 secondes pour remplir les items après avoir lu toute la copie, et au contraire ça facilite l'attribution de la note finale où j'élimine le trop-intuitif. Je ne prétends pas non plus que mon évaluation soit objective, mais il faut assumer la subjectivité de l'évaluation sans lui laisser la possibilité d'être pour autant injuste.