

IUFM d'Aix-Marseille
Formation de formateurs
Journée du 16 mars 2004

***Le moment de l'évaluation, ses objets, ses fonctions :
déplacements, ruptures, refondation***

Yves Chevallard

Le texte ci-après résulte de la mise en forme de notes éparées, réunies à titre de support d'une brève intervention dans le cadre de la journée de travail du 16 mars 2004. Il n'a en conséquence aucune prétention à brosser un tableau complet, même très résumé, de la question abordée.

1. Déplacement 1 : d'un sémantisme profus à une ascèse du sens

1.1. Estimer la *valeur* d'un « objet », ou plutôt lui *attribuer* une valeur (d'une manière ou d'une autre), bref, *évaluer* constitue une activité spécifiquement humaine, qui peut porter sur *n'importe quel* « objet » *o*, être le fait de *n'importe qui*, *z*, se réaliser *dans une institution quelconque I*. Mais ces trois variables (quoi ? qui ? où ?) *ne sont pas indépendantes* : tout « objet » peut être évalué, mais pas dans *n'importe quel* contexte institutionnel ; et, dans une institution donnée, pas par *n'importe qui*.

1.2. Même si la tentation d'évaluer hante souvent l'abord d'un objet *o* par un sujet *z* d'une institution *I*, ce n'est pas là, en droit comme en fait, le seul type de « verdict », de « vrai dit » (*verdictum*, en latin) qui puisse s'énoncer à propos de *o*. Un jugement d'évaluation relève certes de l'activité de *véridiction* à l'endroit de l'objet, mais *l'inverse n'est pas vrai*. Ce qui prépare et fonde, en droit comme en fait, un verdict d'évaluation, ce sont des verdicts d'*analyse* : dans son principe, l'analyse précède l'évaluation, au double sens logique et chronologique. Le jugement évaluatif n'est souvent, au reste, qu'une expression inadéquate, inaboutie (parce que censurée, par exemple) d'une assertion proprement analytique, sans contenu évaluatif.

1.3. L'absence de contenu évaluatif est normalement liée au fait que, sauf lorsqu'elle dégénère en une activité purement *formelle*, arc-boutée sur des critères que le passage du temps de l'institution aura subrepticement naturalisés, l'activité d'évaluation n'est pas gouvernée par une logique intrinsèque. *Fonctionnel*, l'acte évaluatif est fondamentalement *relatif* : car une valeur ne peut être attribuée à un objet qu'à un point de vue déterminé, par rapport à un certain *usage* social, dans le cadre d'un certain *projet* institutionnel. C'est là un fait fondateur, bien attesté d'ailleurs par les dictionnaires. Le mot de valeur, précise ainsi le *Dictionnaire historique de la langue française* (Le Robert, 1993), « s'emploie à propos de la qualité, de l'intérêt d'une chose, d'où *avoir valeur* " être propre à un certain usage " (1080) ». En anglais, semblablement, « *value* » désigne, selon l'auteur du *Penguin Dictionary of Psychology* (Penguin Books, 1985), « *the quality or property of a thing that makes it useful, desired or esteemed* » – le même auteur ajoutant ce commentaire sans ambiguïté : « *Note the pragmatic aspect implied by this definition; the value of a thing is given by its role in a (social) transaction, the thing itself does not possess value.* »

1.4. Le caractère relatif de tout acte d'évaluation explique en partie les restrictions vues plus haut quant à l'écologie des actes évaluatifs. Tout d'abord, toute institution *I* ne connaît pas l'objet *o*. Ensuite, si elle le connaît – s'il est l'un de ses *objets institutionnels* –, elle ne confère

pas nécessairement à toute *position p* en son sein la *légitimité* (institutionnelle) d'évaluer *o*. Enfin, même si les sujets *z* de *I* occupant en son sein une position *p* ont, en tant que tels, à connaître l'objet *o*, le rapport à cet objet que leur impose *I* ne fait pas nécessairement apparaître celui-ci comme partie prenante éventuelle de tel projet ou usage social ; en sorte que, si évaluation il y a, ce ne sera pas forcément dans telle perspective souhaitée par telle personne ou telle institution.

2. Déplacement 2 : la dialectique du silence et de la véridiction

2.1. Les pratiques d'évaluation d'objets *o* en tous genres sont si profuses qu'elles peuvent apparaître comme une gêne, voire comme la source de bien des malheurs ! Certains philosophes antiques ont ainsi fait du *refus de juger*, de prendre parti, le fondement absolu de la *vie bonne*.

2.2. Ainsi en va-t-il, notamment, de Pyrrhon (365-275 av. J.-C.). L'attitude à laquelle porte la doctrine pyrrhonienne est celle de la *suspension de jugement*, l'*epoché*. En pratique, cette absence de jugement constitue le *fond de la vie ordinaire*, sur lequel s'élève le *bruissement des jugements de valeur*. La vie est donc faite *presque toujours de non-jugement* ; mais un jugement *est toujours possible*, et advient de temps à autre – trop souvent, au gré de Pyrrhon (ou de Montaigne, son disciple).

2.3. L'usage extensif, attrape-tout du mot d'évaluation, effet pervers du succès populaire de la notion de « *formative evaluation* » lancée par Michael Scriven en 1967, et qui tend constamment à confondre analyse et évaluation, va de pair avec la prolifération contemporaine des actes d'évaluation (par exemple chez les professeurs) et la pression de la demande de tels actes (par exemple de la part des élèves). De cette évaluation ubiquiste on peut voir une manifestation extrascolaire dans l'engouement des jeunes pour se présenter devant des jurys d'audition (du genre de celui d'émission la « La Nouvelle star », sur la chaîne de télévision M6), alors même qu'ils n'ont encore rien construit – aucun objet *o* – qui puisse être évalué, quel que soit le projet social dans lequel on s'évertue à le situer. En ce cas, l'observance du principe d'*epoché* devrait conduire les jurys ainsi sollicités à *refuser de prononcer quelque verdict que ce soit*, en évitant même les gloses phatiques d'apparence évaluative (« Très bien ! », etc.), qui ne sont que des ponctuations de l'échange, sans contenu d'évaluation. Il est un temps pour faire et pour analyser ce faire, et qui ne doit pas être gaspillé au nom du prurit évaluatif « moderne ». Il est un temps où il n'est pas encore temps d'évaluer ; et, au reste, il est un temps où il n'est plus temps de le faire.

3. Déplacement 3 : les objets à évaluer et les raisons de le faire

3.1. À l'inverse, il y a un *moment de l'évaluation* qui doit chaque fois trouver son temps juste. Dans le faire humain, quel qu'il soit, l'évaluation – que vaut ce que nous avons fait, ce que vous avez fait, et que vaut ce faire même ? – est une scansion essentielle, moment de véridiction qui, malgré les souvenirs d'enfance, n'est nullement une invention de l'École, et participe en vérité de la « respiration » même de toute activité humaine.

3.2. Comment analyser ce faire humain ? On a dit que l'objet *o* évalué l'est par rapport à un certain usage social, à un certain projet institutionnel. Tout « usage » – on l'admettra ici – se ramène à la mise en œuvre d'une *praxéologie*, c'est-à-dire d'un quadruplet $O = [T/\tau/\theta/\Theta]$ formé, dans sa version la plus simple (praxéologie « ponctuelle »), par un *type de tâches* *T*, une *technique* τ , manière d'accomplir les tâches *t* du type *T*, une *technologie* θ , discours

raisonné (*logos*) justifiant et éclairant la technique τ , enfin une *théorie* Θ , discours justifiant et éclairant la technologie θ . L'objet o à évaluer doit donc être « apprécié » en tant que participant de la mise en œuvre d'une ou plusieurs praxéologies O_i . Cet objet o peut être un objet *activé* par la mise en œuvre des O_i . Ce peut être aussi le rapport $R(x, o')$ d'une *personne* x ou le rapport $R(X, o')$ d'un *collectif* X à un certain objet o' activé par les praxéologies O_i , c'est-à-dire un « objet » *activant* la mise en œuvre de ces praxéologies.

3.3. Simplifions un peu. Dans une classe de mathématiques, on travaille à élaborer ou à ré-élaborer, à étendre une certaine praxéologie mathématique O . Ici, en un premier temps, le collectif créateur de la praxéologie, X , c'est-à-dire la classe, est le même que le collectif destinataire de O : la classe crée d'abord pour son usage et celui de ses membres. Cela noté, la classe X doit d'abord – selon une consécution logique plutôt qu'en une roide succession chronologique – faire sien un certain type de problèmes T (comment, par exemple, déterminer le prix hors TVA quand on connaît le prix TVA incluse ?), élaborer une technique τ relative à T , une technologie θ de τ , etc. Tel est son faire. Celui-ci doit régulièrement être observé, en des prises d'information qui distinguent sans les dissocier le processus de création et ses produits, puis analysé – à ce double titre : comme processus, et à travers ses produits. Telle est la première urgence, hors de laquelle rien ne vient à l'existence. Au cours de ce processus de création praxéologique, le silence évaluatif règne. Des rapports personnels se créent aux objets o qui, peut-être, constitueront O , ou que la création de O mobilise ici et maintenant. Certains de ces objets auront une existence éphémère et disparaîtront sans laisser de traces. Mais d'autres objets vont se stabiliser, auxquels, dans cette institution qu'est la classe au travail, des rapports *institutionnels* vont émerger. C'est l'existence de tels rapports institutionnels transcendants aux personnes qui fonde en raison le projet d'évaluer les rapports *personnels* des élèves en les référant à la *norme* que le travail d'institutionnalisation aura en quelque sorte hypostasiée.

3.4. On l'a suggéré : l'opération d'évaluation doit être entendue aussi en un sens plus large que ce sens tout classique. Derrière l'évaluation des rapports *personnels* (qui, même en droit, n'est pas évaluation des *personnes*), se profile l'évaluation de la *norme elle-même*, soit du rapport institutionnel qui sert d'étalon. Que vaut, en fait, l'organisation mathématique qui s'est construite et a été institutionnalisée ? Au-delà de l'interrogation sur la maîtrise, par telle personne ou tel collectif, de telle technique ou de sa technologie, on trouve alors l'interrogation *sur la technique elle-même* – est-elle puissante, maniable, sûre, robuste aussi ? Cette évaluation – à laquelle les usages scolaires font, il est vrai, une part fort petite – est ici formatrice, on l'a souligné, non d'une *personne*, mais d'une *praxéologie*. Réformatrice, elle relancera l'étude, suscitera la reprise de tel ou tel moment du travail, voire de l'ensemble du parcours de création praxéologique.

3.5. On touche ici à l'une des grandes questions que soulève l'activité *générique* d'évaluation : celle de ses *raisons d'être*. Pourquoi évalue-t-on ? À quoi cela sert-il ? Quelles sont les *fonctions* de l'évaluation ? D'une manière générique, les jugements de valeur sont *l'antidote à l'indifférence*, au « tout se vaut », au « pas plus ceci que cela » pyrrhonien. L'indifférence justifie de ne (plus) rien faire. Les verdicts d'évaluation combattent donc les effets de l'indifférence en justifiant, voire en impulsant *l'action visant le changement* (vers le « mieux »). Ils luttent contre l'inertie liée à *l'arbitraire des manières de faire*, à propos de laquelle l'historien Paul Veyne a pu écrire (dans sa préface au livre de Peter Brown, *Genèse de l'Antiquité tardive*, Gallimard, 1983) : « ... l'homme est un être indéterminé, inachevé ; rien ne lui indique impérativement ce qu'il doit faire, dire et croire ; il n'a ni nature, ni infrastructure, ni raison qui ne soit vulgaire rationalisation du monde humain comme il va. Il

ne sait pas comment se tenir. Alors ce dormeur éveillé se décide à prendre une posture et, tant qu'il la conserve, il n'en connaît pas d'autre... » L'évaluation est donc un *opérateur de changement de « posture »*.

3.6. Cependant, cet opérateur oblige à subsumer – au moins localement – la *variété* (la pluralité) sous une (supposée) *relation d'ordre* sur l'ensemble des états possibles. Pour cette raison – et tel est le fond de la critique pyrrhonienne –, elle nous met en danger de prétendre *faire mieux*, ou *avoir mieux*, dans une quête indéfinie du meilleur et du mieux. De là qu'il soit nécessaire de rechercher un *bon usage* de l'évaluation, qui incite au changement souhaité, non à la quête sans fin.

4. Déplacement 4 : déverrouiller les pratiques ordinaires d'évaluation scolaire

4.1. Un danger solidaire guette l'évaluation, que l'expérience de l'évaluation scolaire, resserrée autour de la *notation*, révèle : le danger d'une pratique lovée sur elle-même, devenue autonome, regardée de fait comme auto-justifiée, porteuse d'une signification et d'une valeur supposées *intrinsèques*. Dans l'évaluation scolaire, deux universaux anthropologiques s'entrecroisent : la dialectique du silence et de la véridiction cernée plus haut tente de s'y articuler à l'intérieur d'une dialectique *du don et du contre-don*. Ce dernier aspect est déterminant. Le professeur fait don d'un cours et offre à la classe un sujet de travail. L'élève lui remet sa copie. Le professeur répond à son tour : pour nombre d'élèves, aujourd'hui, la *note* qu'il assigne le libère pour l'essentiel de la dette contractée envers l'élève au moment où il en a accepté la copie. La dialectique du don et du contre-don, ici, tourne court : le marché se conclut trop vite, au prix d'un peu de verroterie numérique que chacun feint de regarder comme réglant la transaction adéquatement. Casser cette mécanique est essentiel.

4.2. La contre-prestation du professeur, il est vrai, ne se limite pas toujours à cela : en nombre de disciplines, par delà l'attribution de la note et la restitution de la copie « annotée », le contre-don professoral comporte une « correction » que les annotations anticipent et que les élèves recevront souvent sans enthousiasme, suivie elle-même de la remise d'un « corrigé » pour solde de tout compte. Mais il faut interroger cette notion même de correction. Le vocabulaire de l'école est un vieux vocabulaire, qui a beaucoup travaillé, qui s'est beaucoup distendu. Le latin *corrigere* signifie « redresser ». Derrière le vocabulaire de la correction, il y a une construction praxéologique ancienne, qui se survit dans les pratiques de correction actuelles, et qu'on peut résumer ainsi : l'élève doit être corrigé – au sens habituel du terme ! Il a une mauvaise posture, une mauvaise manière de faire, il doit être redressé. Dans cette problématique correctrice, le problème n'est pas tant de *créer*, de construire une manière de faire, que de corriger la manière de faire supposée *déjà là*, inscrite dans les corps et les esprits. Une certaine tournure a été prise, qu'il convient de policer, de réformer, de civiliser. Tout se passe, curieusement, comme s'il y avait un donné quasi naturel, et qui soit vicieux. La correction consistera donc à dire ce qui ne va pas, à morigéner le fautif, etc.

4.3. Comment peut-on penser qu'un jeune enfant, qui ne sait pas faire la division, sait la faire mais la fait mal ? Si j'apprends à faire quelque chose, ou à *penser* quelque chose, *a priori* je ne sais pas le faire, je ne sais pas le penser : il n'y a encore rien à corriger. Corriger la manière de marcher d'un jeune enfant a un sens ; d'un très jeune enfant, non. Car cela supposerait qu'il sache marcher, et que sa manière de marcher soit vicieuse, imparfaite, défectueuse. Mais comment peut-on prétendre corriger sa manière de faire et de penser quelque chose qu'il ne sait ni faire ni penser – la division par exemple ? Ainsi s'opposent deux problématiques de l'instruction, de l'éducation, de la formation. Dans la problématique « moderne », on suppose

au contraire quelque chose qui se construit *hic et nunc*, qui n'était pas là, qui n'est pas un donné pur de la société, sans parler même de la nature ! Cette chose qui se construit, bien entendu, on peut toujours dire qu'elle se construit de manière vicieuse, et que, au fur et à mesure qu'elle se construit, il faut la corriger. Mais on peut penser aussi que cet argument-là est le fruit opportuniste, tardif, d'une tentative spontanée d'adapter l'ancien (la correction d'un préformé supposé) au nouveau, la création d'une réalité sinon indéterminée, du moins *sous-déterminée*.

4.4. La configuration idéologique en jeu dans le paradigme de la « correction » s'est sans doute transformée depuis le XIX^e siècle : elle n'en travaille pas moins, sourdement, les pensées et les pratiques contemporaines de l'évaluation. Dans la pureté de son principe, le paradigme ancien s'énonce ainsi : « Je prends ce que vous avez fait et je le regarde comme quelque chose qui est imparfait, qu'il faut corriger, etc. » Dans une forme modernisée, on se livre avec empressement, parfois avec passion à l'*analyse des erreurs*, forme sublimée et souvent impitoyable de la correction à l'ancienne. Avec cela, pourtant, on est fort loin des pratiques authentiques de la connaissance ! On sait par exemple que, au grand dam de régents épris de correction, et qui voudraient tant que l'histoire s'écrive à l'encre rouge, dans l'histoire des mathématiques l'erreur est promptement oubliée – parce qu'on ne la refera plus, parce que ce n'est plus comme ça qu'on fait désormais, parce que des chemins ont été ouverts qui rendent impossibles les anciennes errances...

5. Déplacement 5 : pour une refondation des praxéologies d'évaluation

5.1. Comment penser l'essentiel du jeu – scolaire ou extrascolaire – avec la connaissance, sur lequel l'acte évaluatif va imprimer sa marque dynamisante ? Le schéma de départ est le suivant. Sous la direction d'un ou de plusieurs directeurs d'étude (dont on note l'ensemble Y), un collectif X tente d'apporter une réponse R à une question Q – une question qui peut être par exemple « comment déterminer le prix hors TVA quand on connaît le prix TVA incluse ? » ou mille autres encore, grandes et petites. X , Y et Q forment ensemble un *système didactique*, $S(X, Y, Q)$; le fonctionnement de ce système didactique doit produire une réponse R , ce qu'on peut noter ainsi : $S(X, Y, Q) \rightsquigarrow R$.

5.2. Que fait le collectif X pour produire une réponse R ? R est en règle générale une praxéologie $[T/\tau/\theta/\Theta]$, ou du moins participe de la création d'une telle praxéologie : au plus large, la création de R suppose donc que, à travers une pluralité de *moments didactiques* qu'on ne détaillera pas ici, on construise T , τ , θ , Θ . De quels matériaux s'alimentent cette étude ? Un schéma de réponse peut être apporté dans les termes suivants, qui explicite le cadre fondateur de toute praxéologie d'étude et de recherche qu'est la *dialectique des médias et des milieux*¹. Un premier geste didactique consiste à *observer* les réponses R^\diamond déposées dans la culture et mises en avant dans les *médias* accessibles. Il faut ensuite *analyser* ces réponses « estampillées » R^\diamond : là se noue véritablement la dialectique entre médias et milieux, ces derniers fournissant les pierres de touche – théoriques et expérimentales notamment – du travail d'analyse. Arrive alors le temps d'*évaluer* les réponses R^\diamond ainsi analysées, geste qui conduira peut-être à réaliser de nouvelles observations ou à pousser plus avant certaines

¹ On entend ici par *média* tout système de mise en représentation du monde à l'adresse d'un certain public. Le cours du professeur de mathématiques, le prêche d'un clerc, le journal d'une présentateur de télévision relèvent en ce sens du système des médias au sens large. Un *milieu* est un système qu'on peut regarder comme dénué d'intention « didactique » dans la réponse qu'il peut apporter, de manière explicite ou implicite (il faut alors interpréter son « comportement de réponse »), à telle question déterminée : il se comporte à *cet égard* comme un fragment de « nature ».

analyses. Le *projet* qui fournit ses critères à l'acte évaluatif est ici tout net : quelle est la *valeur* de ces réponses en tant que matériaux de construction d'une réponse R – qu'on notera R^\heartsuit – diffusable au sein du collectif X et défendable devant ce collectif ? Les gestes à accomplir à partir de là, au-delà de cette césure évaluative, sont ceux-là mêmes que l'évaluation anticipe et sur lesquels elle s'oriente : *développer* une réponse R^\heartsuit , *diffuser & défendre* la réponse R^\heartsuit ainsi produite.

5.3. Que devient ici le rite de la « correction » ? Le professeur a offert à la classe une question Q et une partie des matériaux utiles² pour apporter à Q une réponse R . Ensemble, sous la direction du professeur, on a examiné des réponses R^\diamond : on les a observées, analysées, évaluées ? En autonomie didactique relative, dans un « devoir à la maison » par exemple, on a ensuite développé *sa* réponse R^\heartsuit , on l'a « publiée » et défendue dans la copie – individuelle ou collective – qu'on a remise au professeur. Que vaut R^\heartsuit regardée désormais, par la classe X , comme une « simple » réponse R^\diamond parmi d'autres ? Question décisive ! La valeur de cette production individuelle ou collective tient *dans ce qu'elle apporte* au projet didactique de la classe : la construction d'une réponse R^\heartsuit commune, partageable. Cela conduit à porter son regard, en premier lieu, non sur la « réunion » des échecs des élèves, mais sur la réunion *de leurs réussites*. À partir des différentes réponses R^\diamond (dont plusieurs sont en partie ou en totalité erronées) qui ont été apportées par les différentes équipes en lesquelles la classe s'est scindée en cette occasion, la classe unie développera, sous la direction du professeur, une réponse commune, qui sera consignée par écrit – c'est le « corrigé », à fonction de bilan d'étape –, avant de donner lieu à une institutionnalisation en bonne et due forme.

5.4. Dans la perspective praxéologique précédente, que devient le thème de l'erreur ? Ce que l'on nomme ainsi va retrouver son sens premier : on regardera l'erreur *d'abord* comme un symptôme de l'effort pour créer une technique (ou la technologie qui doit l'accompagner), mais d'un effort *qui, jusqu'ici, n'a pas abouti*, chemin d'errance qui n'a, jusqu'à présent, conduit nulle part. Que faire de l'erreur, de l'errance ? Il faut là suspendre son jugement, et revenir au labeur d'observation, d'analyse, de développement ! Dans le cadre du projet collectif où s'inscrivent les « œuvres » des élèves, on se demandera ainsi *comment* la tentative inaboutie aurait pu aboutir (ce qui exigera souvent de se demander *pourquoi* elle n'a pas abouti, sur quoi elle a *buté*). Prenons un exemple rapide, à valeur essentiellement illustrative. Supposons un élève qui, familier des équations du *premier* degré, rencontre pour la première fois une équation du *second* degré, par exemple $x^2 - 3x + 2 = 0$. Mû par un *habitus* anciennement constitué, il récrit l'égalité précédente sous la forme $(x - 3)x = -2$, puis en tire la « valeur » $x = 2/(3 - x)$. Las ! cette valeur de x dépend... de x . Erreur ! Certes. Mais quelle est cette ligne d'erre sur laquelle l'élève a fait ainsi ses premiers pas ? Où conduit-elle ? Nulle part, dira-t-on trop vite. À un abord *subjectiviste* de l'œuvre inachevée de l'élève (dont lui-même ne reconnaîtra pas d'abord l'inachèvement), il convient alors de substituer une analyse *objectivante*, qui tente d'y saisir l'amorce d'une technique possible, peut-être hors de portée, mais qu'il n'est pas interdit, qu'il est souvent utile d'explorer³, dans la visée même (ici) d'une technique purement algébrique...

² La dialectique des médias et des milieux est, dans la tradition scolaire, contrôlée de façon étroite, voire obsessionnelle. Resserrée autour du « cours » du professeur, elle tend en conséquence à se réduire au minimum dogmatiquement nécessaire. Ce qui suit fait volontairement référence à une dialectique ouverte, qui reste en grande partie à construire au plan institutionnel.

³ Il y a là en effet l'ébauche d'une des techniques les plus précieuses qui soient : une technique de *point fixe*. En partant de $x_0 = 0$, on obtient ici, à l'aide d'une calculatrice, $x_1 = 2/(3 - x_0) = 0,6666666...$, $x_2 = 2/(3 - x_1) = 0,8571428...$, $x_3 = 0,9333333...$, ..., $x_{20} = 0,9999995...$, ..., $x_{25} = 0,9999999...$, ..., $x_{27} = 1$, $x_{28} = 1$. Ainsi voit-on apparaître, calculatrice aidant, une des solutions de l'équation, $x = 1$, qu'il restera à ressaisir algébriquement.