

Vers une approche systémique de l'évaluation ingénierie de l'évaluation à partir du modèle de Stufflebeam

Auteur : Abdennasser Naji

Professeur agrégé et inspecteur pédagogique

Résumé

La problématique de l'évaluation est multidimensionnelle et de ce fait elle doit être approchée d'une manière systémique et globale. Le système éducatif est un tout indissociable qui englobe des inputs, des processus et des outputs en interactions dynamiques qui ne peuvent être évalués séparément au risque d'engendrer des décisions erronées : l'évaluation étant l'activité aidant à prendre des décisions. A ce titre, elle doit moins se focaliser sur les performances que sur les processus éducatifs de toutes sortes (de réalisation, de soutien et de management) et à tous les niveaux. Evaluer les performances ne serait plus qu'un instrument au service de l'amélioration continue de la qualité de l'éducation.

Mots-clés

Evaluation – systémique – curriculum – conception – mise en œuvre – processus – indicateurs - qualification

Introduction

Le concept de validité systémique de l'évaluation¹ nous interpelle sur la nécessité d'opter pour une approche globale et systémique de l'évaluation en mesure d'engendrer une transformation des pratiques, qui donne lieu à des changements utiles à tous les niveaux du système éducatif: enseignants, enseignés et gestionnaires. Cette validité est d'autant plus élevée que l'évaluation est approchée sous ses différents aspects : cognitif, sociologique, didactique, etc. et que le sujet évalué est vu non comme un élément isolé mais comme une composante faisant partie d'un système global, et de ce fait il serait indispensable de prendre compte de ses interactions avec les autres composantes. Le modèle de Stufflebeam² constitué des 4 piliers ; contexte, intrants, processus et produit ; représente un bon guide pour l'évaluation systémique. Par ailleurs, chaque sujet d'évaluation (individu ou organisation) est le résultat d'un certain nombre de processus ayant un cycle de vie comprenant deux étapes fondamentales : la conception et la mise en œuvre. Le processus d'enseignement/apprentissage s'appuie sur plusieurs éléments dont le plus important est le curriculum. Celui-ci, avant d'être mis en œuvre passe par l'étape de conception. Pour nous aider à prendre les décisions pertinentes, on ne pourrait faire l'économie d'une évaluation au niveau de la conception. A défaut, on risquerait d'entériner un curriculum inadéquat. La suite de l'article donne un aperçu sur quand et comment évaluer un système éducatif et ses différentes composantes.

¹ (Frederiksen, J.-R. & Collins, A. (1989). A Systems Approach to Educational Testing. Educational Researcher, 18, 27-32.

² Stufflebeam, D. et al. (1980). L'évaluation en éducation et la prise de décision. Ottawa : NHP.

Évaluation des solutions

A la fin de l'étape de recherche de solutions curriculaires possibles et après avoir étudié l'existant, on se retrouve avec une série de propositions parmi lesquelles il va falloir choisir celle qui convient le mieux. C'est l'objet de l'étape d'étude et d'évaluation des solutions. Des outils permettent de choisir la solution la mieux adaptée vis à vis des critères que l'on s'est fixés, et c'est justement le choix de ces derniers qui constitue la difficulté majeure dans cette opération de tri des solutions. Toutes celles qui ont résulté du travail de conception doivent nous garantir en principe un système éducatif qui répond aux besoins que l'on compte satisfaire et même si, malgré les précautions prises, il y en a qui n'offrent pas cette garantie, on les éliminera lorsque le travail de validation des solutions aura été fait. Il y a donc lieu de penser à d'autres indicateurs que l'adéquation des solutions avec les besoins tels la rentabilité, le coût, l'efficacité ou le délai de mise en application pour départager les solutions en compétition. Ces indicateurs ou d'autres sont déterminants dans le choix de la (ou les) solution (s) qui sera par la suite expérimentée puis évaluée avant d'être revue et corrigée s'il y a lieu, pour enfin être définitivement adoptée.

Qualification du curriculum

Qualification interne et qualification externe

Dans une première approche de la problématique de la qualification on propose de la décomposer en une qualification interne réalisable au sein même du système éducatif et une qualification externe s'effectuant en dehors de ce système.

Chaque étape importante de la conception doit être validée avant de passer à l'étape suivante. Lorsque c'est fait avec succès, une qualification des choix définitifs est nécessaire. Les modalités d'application doivent être clairement définies ainsi que l'attribution des responsabilités au sein de l'équipe chargée de cette mission.

Des documents supports sont à prévoir pour décrire sommairement le déroulement des actions de qualification et notifier les résultats auxquels elle a aboutis.

Qualification matérielle et qualification immatérielle

Le système éducatif est un grand ensemble qui regroupe plusieurs sous-ensembles dont chacun peut à son tour être décomposé en plusieurs éléments. L'organigramme technique permet de visualiser l'arborescence de ce système et d'illustrer bien sa complexité. Il va sans dire que la conception dont on parle doit prendre en compte cet aspect des choses sans pour autant occulter la vision systémique selon laquelle on doit considérer le système éducatif dans sa globalité et coordonner les tâches entre les différentes équipes qui s'occupent de ses sous-ensembles et éléments afin de garantir la cohésion et la cohérence de l'ensemble. Celui-ci est formé de composantes matérielles tels les outils didactiques et les livres scolaires, et d'autres immatérielles comme les savoirs et les méthodes pédagogiques. La nature de la qualification à appliquer diffère selon que l'on a affaire à l'une ou à l'autre composante.

Pour les objets, la qualification que l'on va nommer désormais "qualification matérielle" peut se faire en réalisant des prototypes qui seront soumis à l'expérimentation afin de juger de la validité ou non des choix effectués lors de la conception. Ceci est aussi valable pour le livre scolaire dont une première version peut être expérimentée avant le début de l'année scolaire sur un groupe restreint d'élèves en choisissant un échantillon de quelques leçons sur lesquels les enseignants vont tester la validité du livre dans les situations réelles d'enseignement. Ce premier test étant valable, le livre scolaire peut être autorisé à l'usage en considérant la première année d'utilisation comme une période de qualification à l'issue de laquelle on fera une évaluation pour décider de l'adoption définitive du livre.

Pour les éléments immatériels, le processus de qualification que l'on nommera dorénavant "qualification immatérielle" est plus compliqué ; ni prototypes ni prélèvement d'échantillons ne sont possibles dans ce cas. La seule chose faisable c'est de passer directement à l'exécution mais en veillant toutefois à rapprocher autant que faire se peut les périodes d'évaluation afin de remédier à temps à tout dérapage ou anomalie. Il est certes possible d'expérimenter les solutions de conception sur un échantillon-pilote avant de décider de leur généralisation, mais vue la vitesse vertigineuse à laquelle tourne le monde aujourd'hui il devient très risqué d'agir de la sorte. En effet, ce n'est pas exagéré d'affirmer qu'en adoptant cette voie on peut se retrouver dans la délicate situation suivante : on n'a pas encore achevé la qualification d'une solution qu'on est contraint de la

substituer par une autre tellement les besoins changent rapidement. C'est dire la nécessité de réfléchir à des modalités de qualification qui puissent s'exécuter avec le décalage le plus réduit possible sur l'application effective des solutions, voire en parallèle avec celle-ci.

Théoriquement c'est faisable, cependant en pratique c'est difficile mais ce n'est pas impossible pour peu que les modalités d'application soient définies avec le plus grand soin. Il faut penser notamment à la manière de remplacer une solution jugée invalide par une autre mieux adaptée, et aux perturbations que cela peut générer. Car le problème qui se pose avec acuité c'est celui de la durée d'exécution d'une solution ayant trait au système éducatif et l'impossibilité pratique de l'interrompre avant son terme. On ne peut évidemment pas décider de suspendre une formation ou de l'annuler au milieu de son déroulement sous prétexte que la qualification n'a pas donné de résultats probants, vues les nuisances morales et matérielles que cela risque d'engendrer chez les personnes qui subissaient cette formation. Tout ce qu'on peut faire dans une telle situation c'est de tenter dans la mesure du possible d'apporter les remèdes nécessaires aux problèmes révélés par la phase de qualification et si celle-ci conclue à l'invalidité totale de la solution envisagée un comité ad hoc sera constitué d'urgence pour prendre la décision qui s'impose. Ce comité ne décidera de remplacer la solution invalide qu'après un examen approfondi de l'impact tant psychologique que socio-économique d'une telle décision sur les principaux concernés. Il doit aussi élaborer un plan de réinsertion sociale pour les gens dont l'avenir professionnel serait menacé.

Rappelons une fois encore qu'agir de la sorte n'émane pas d'un choix délibéré mais résulte d'une obligation. C'est bien évidemment le risque de déphasage entre les attentes du marché et les profils des sortants du processus d'enseignement qui engendre cette contrainte. C'est un risque tellement énorme qu'on n'a plus aucune autre alternative pour la qualification. Il reste à déterminer les critères par rapport auxquels il sera décidé du degré de réussite de cette dernière. Il y en a certes plusieurs, mais ils ont tous un seul dénominateur commun : la satisfaction des besoins des utilisateurs.

Des règles à respecter

Chaque fois que cette thèse n'est pas confirmée (pas à 100% bien sûr mais avec une marge d'erreur à définir), la qualification sera déclarée non concluante. Un rapport sera ainsi rédigé par l'équipe responsable de la qualification pour mentionner les raisons de l'échec de celle-ci et proposer les amendements qu'elle juge adéquats. Un comité spécial se basera sur le contenu de ce rapport pour sanctionner la qualification par une décision finale qui s'articulera autour de l'un ou de l'autre des axes suivants :

- perfectionner la solution objet de la qualification en remédiant aux problèmes constatés ;
- remplacer cette solution par une autre plus performante en prenant en compte les propositions faites par les responsables de la qualification.

Quelle que soit la nature de la décision prise, elle sera communiquée à toutes les parties concernées.

Chaque élément qualifié doit être accompagné d'une fiche permettant de connaître son état « qualificatoire » en précisant entre autres le nom de l'organisme qualificateur, la date, le coût et les conclusions de la qualification.

Evaluation des intrants

Quelque soit le système adopté pour l'orientation des élèves, une évaluation à l'entrée de chaque nouveau cycle de l'enseignement est nécessaire. Il a le rôle d'identifier les élèves qui possèdent le profil adéquat et de détecter ceux qui présentent des anomalies. Les premiers sont évidemment acceptés, les seconds sont mis en quarantaine jusqu'à la prise de l'une des deux décisions suivantes :

- leur barrer carrément la route et les aiguiller vers d'autres établissements ou d'autres filières plus adaptés à leurs aptitudes ;
- les accepter mais sous réserve de leur faire subir un programme spécial de remise à niveau, afin de remédier aux lacunes détectées lors de l'évaluation.

L'évaluation des candidats se fait en considération de critères relatifs aux compétences requises par l'établissement d'accueil et en fixant les niveaux de performances exigés.

Les résultats des différentes évaluations sont enregistrés et conservés dans des fiches classées selon les établissements de provenance des élèves. Ils serviront par la suite à évaluer ces établissements afin de décider de l'éventuelle implication de ces derniers dans une relation de partenariat.

L'identification de l'établissement d'origine pour chaque élève est nécessaire pour déterminer la part de responsabilité de chacun dans un éventuel échec de celui-ci.

Des fiches sont élaborées à cet effet. Elles définissent l'itinéraire scolaire de l'élève avec des précisions sur les établissements, les enseignants, les compétences acquises, les niveaux de performances, etc.

Procédures, instructions et enregistrements

Pour des fins de traçabilité, chaque élève doit être identifié par une fiche donnant toutes les informations utiles permettant de décrire son historique. C'est une sorte de portfolio regroupant les travaux de l'élève, l'état des différentes évaluations et les progrès réalisés.

Des documents décrivant les procédures de réalisation de chaque activité doivent être élaborés. Ils doivent pouvoir donner une réponse univoque à des questions telles que :

- quels moyens sont les plus adaptés pour satisfaire les exigences d'un programme d'études ?
- quelle voie pédagogique emprunter pour faire acquérir aux élèves une compétence donnée ?
- comment préparer la fabrication des matériels pédagogiques ?
- quels moyens et quels types d'évaluation utiliser, et à quels moments du déroulement du processus de l'enseignement faut-il les employer ?
- combien de temps faut-il consacrer à l'acquisition de chaque compétence ?

Afin d'aider les différents acteurs à s'acquitter convenablement de leur tâche, et pour assurer une certaine régularité dans l'exécution des activités éducatives, des instructions de travail sont à élaborer. Ce sont généralement des fiches d'ordres divers. On peut citer, par exemple, la grille d'observation de la classe pour les inspecteurs, la fiche de préparation des leçons et la grille d'évaluation des compétences pour les enseignants, la carte des compétences pour l'administration de l'école, la check-list pour l'évaluation de la compatibilité des contenus des leçons avec les programmes, les modes d'emploi des matériels didactiques, etc.

Les matériels didactiques sont codifiés et des fiches signalétiques indiquant notamment dans quel chapitre du programme ils doivent être employés sont élaborés.

L'état de l'évaluation des produits

L'évaluation est un paramètre clé dans la démarche qualité, non pas pour ses effets dissuasifs ou son pouvoir de sanction, mais surtout pour sa dimension informative. Il permet de renseigner sur la situation réelle du produit. En la comparant à la situation attendue on tire les enseignements qui s'imposent, ce qui permet d'agir en connaissance de cause sur le processus, de sorte qu'il soit toujours apte à produire la qualité requise. On fait alors de la régulation, didactique pour l'élève qui est renseigné sur ce qu'il doit apprendre et managérial pour l'organisation qui est renseigné sur ce qu'elle doit améliorer. L'évaluation joue ainsi un rôle dynamique non pour prendre la photo d'une situation mais pour la faire évoluer et progresser. Elle est formative par nature, et certificative par déformation³, et même lorsqu'elle se fait dans un but sommatif, il faut toujours exploiter ses résultats pour améliorer le processus de l'enseignement. L'évaluation reste sans doute une source intarissable de perfectionnement, car il permet de donner l'image réelle du processus à un instant donné. Pour qu'elle s'acquitte avec brio de cette tâche difficile, il faut respecter au moins deux conditions :

- augmenter autant que faire se peut la fréquence d'évaluation afin de détecter aussitôt que possible toute défaillance, et avoir suffisamment de temps pour la corriger avant qu'il ne soit trop tard ;
- libérer les élèves de toutes les craintes et contraintes inhérentes aux examens, ce qui pourrait les inciter à aborder l'opération de contrôle avec plus de sûreté et de confiance en soi d'une part, et réduire les tentatives de fraude et de tricherie d'autre part.

La satisfaction de ces deux conditions est susceptible de rendre à l'évaluation sa vraie raison d'être : assurer un retour d'information qui peut contribuer à corriger les erreurs d'apprentissage. Chaque élève doit avoir sa propre

³ Cardinet, J. et Laveault, D. (1996). Dix années de travaux européens et nord américains sur l'évaluation : quelles lignes de force ? *Mesure et Evaluation*, 18, 1-25.

fiche d'évaluation qui comprend en plus des notes résultant de l'évaluation de chaque compétence, un recensement des difficultés rencontrées, les mesures prises pour les surmonter et les résultats obtenus par l'élève après l'exécution des actions correctives.

Evaluation des produits

Le produit fini ne peut être accepté par le client que s'il passe l'épreuve du contrôle final avec succès. Celui-ci répond à un certain nombre de règles établies en commun accord avec le client. En principe, on a donc le fournisseur d'un côté, le client de l'autre, qui définissent conjointement les spécifications du produit qui feront impérativement l'objet du contrôle final. C'est un schéma difficile à calquer avec rigueur et fidélité dans le cas de l'enseignement pour trois raisons qui semblent évidentes :

Le client n'est pas une entité clairement matérialisée, avec qui le fournisseur peut négocier directement ou par le biais d'intermédiaires. C'est plutôt une cible mobile qui ne cesse de se mouvoir dans le temps en changeant continuellement d'aspect. De plus, elle ne dispose pas d'interlocuteur officiellement reconnu et autorisé à parler en son nom et à prendre des décisions qui l'engage. C'est un scénario sombre qui ne semble augurer rien de bon en ce qui concerne la qualité du produit. Tous les ingrédients nécessaires pour obtenir celle-ci sont en effet absents : la cible n'est pas parfaitement identifiée, les débouchés ne sont pas assurés et en conséquence le produit est mal ou incomplètement défini. Heureusement que cette situation n'est pas irrémédiable pour peu que les principaux intéressés, les opérateurs économiques en premier lieu, veuillent bien s'impliquer dans la construction du système éducatif et prêter main forte aux acteurs de celui-ci afin qu'ils parviennent à lever tous les défis qui se présentent à eux ;

Le terme contrôle final peut prêter à plusieurs interprétations si nous ne précisons pas la signification que nous lui attribuons ici. Il est en effet, tout à fait légitime d'assigner le terme en question à tout examen prenant la forme d'évaluation sommative, comme on est en droit de réserver son utilisation aux situations d'évaluation qui prévalent à l'interface entre deux classes de niveaux différents. On peut enfin le dédier exclusivement aux examens de sortie qui permettent de porter le jugement final sur l'aptitude du candidat à répondre d'une manière satisfaisante aux attentes du métier qu'il aurait choisi. C'est cette dernière acception que nous retenons pour le terme contrôle final car c'est la seule qui fait référence à l'ultime client du système éducatif : le marché de l'emploi en général. Les autres significations sont toutefois acceptables lorsqu'on se met dans les situations de client-fournisseur que l'on est en mesure de rencontrer dans les différentes étapes du processus de l'enseignement. Chaque situation est un cas à part dont il faut étudier de plus près toutes les spécificités afin d'en maîtriser les différentes facettes de son évaluation.

Les spécifications devant faire l'objet de la vérification finale ne sont pas faciles à circonscrire. S'il est vrai que chaque métier requiert un certain nombre de compétences qui sont parfaitement cernées et complètement définies, il n'en demeure pas moins vrai que la question de savoir quelles caractéristiques les postulants à un métier donné doivent posséder pour être jugés aptes, reste difficile à trancher et ce pour deux raisons, au moins :

- les différences structurelles entre le monde du travail d'un côté (services et fonctions) et le monde de l'éducation de l'autre (matières et disciplines) font qu'il existe rarement une équivalence parfaite entre les compétences exigées par le premier et celles bâties par le second. Chaque fonction exige des capacités qui sont en fait des compétences éparpillées en général dans plusieurs disciplines, et la difficulté est de réussir à déterminer les compétences qui doivent faire l'objet de l'évaluation finale. Tout effort qui tentera de reconstituer, à travers l'examen de sortie, la mosaïque des capacités par les compétences adéquates ne sera que peine perdue en raison de leur diversité et de leur multitude. Ce qu'il est possible de faire pour sortir de cette impasse, c'est de fixer une échelle de priorité sur laquelle il faut se baser pour choisir les compétences à privilégier par rapport aux autres ;
- les compétences terminales faisant l'objet du contrôle final sont en général l'aboutissement d'un processus de cumul d'un certain nombre de compétences intermédiaires, acquises à des étapes différentes de l'enseignement et dans des situations variées qui ont en principe fait l'objet d'évaluations antérieures. L'accomplissement effectif de ces dernières ainsi que leur fiabilité sont deux facteurs essentiels qui, s'ils sont assurés, préjugent de l'acquisition de la compétence fédératrice dont elles font partie, mais sans que cela puisse nous offrir des renseignements précis sur la qualité de cette acquisition. Encore faut-il inventer la formule capable de formaliser avec une certaine précision la relation qui existe entre chaque compétence terminale et les compétences intermédiaires qui lui donnent naissance.

L'approche par compétences tente de combler ces lacunes en conférant à l'école le rôle de former des gens non uniquement « savants » mais aussi et surtout compétents au sens de capables d'agir avec efficacité face à une famille de situations qu'ils arrivent à maîtriser parce qu'ils disposent à la fois des connaissances nécessaires et de la capacité de les mobiliser à bon escient, en temps opportun, pour identifier et résoudre de vrais problèmes⁴.

Documents d'évaluation

Afin que l'évaluation réponde convenablement à sa mission il serait utile de rédiger des documents précisant entre autres les paramètres à évaluer, les moyens recommandés pour le faire, les indicateurs et les fréquences de vérification.

C'est un travail d'équipe qui demande la participation de toutes les parties prenantes afin d'éviter toutes sortes de redondance dans les évaluations pratiquées et pour assurer le maximum de coordination entre les différents acteurs.

Une fois les documents rédigés et approuvés par les responsables il ne faut surtout pas lésiner sur les moyens pour les diffuser et les faire parvenir aux endroits et aux personnes auxquels ils sont prédestinés.

En outre, il faut prendre toutes les dispositions pour assurer une évaluation efficace du produit. Pour cela il serait utile de prévoir un plan précisant surtout :

- les opérations à effectuer et les paramètres à mesurer ;
- les opérations à effectuer concernant les produits de provenance extérieure et les élèves à l'entrée de chaque nouveau cycle ;
- les critères d'acceptation ou de refus ;
- les moyens à mettre en œuvre ;
- les références des normes ou standards à utiliser ;
- les conditions d'environnement à respecter;
- les méthodes statistiques mises en œuvre ;
- les opérations d'évaluation concernant le produit fini.

L'enregistrement des résultats des différentes évaluations mérite un soin particulier car il permet de constituer une base de données que l'on peut consulter à chaque fois que le besoin se fait sentir. Différentes utilisations sont envisageables dont on peut citer à titre indicatif : suivre l'évolution de caractéristiques données, évaluer la qualité des fournisseurs, chercher les causes des anomalies constatées ou analyser une situation donnée. Il serait en outre très utile de faciliter la diffusion de ces enregistrements aux instances et personnes concernées.

Synoptique d'évaluation

Ce synoptique est une représentation graphique de l'insertion des activités évaluatives dans le processus des activités considérées : L'importance cruciale que revêt l'évaluation dans le système éducatif exige que l'on accorde une attention toute particulière à la planification et à l'organisation de cette fonction. Il est ainsi recommandé de définir les différentes articulations de l'évaluation avec les autres activités et fonctions qui jalonnent le processus de l'enseignement, ce qui pourrait être illustré avec clarté par le truchement d'un synoptique.

Une esquisse de celui-ci peut très bien s'établir dès la phase de conception du système éducatif tout en laissant la latitude aux différents acteurs pour l'affiner chacun en fonction de ses besoins propres et des contraintes que la pratique ne manque pas de faire jaillir.

Il est toutefois indispensable que cette esquisse constitue la toile de fond commune sur laquelle toutes les parties sont tenues de s'appuyer pour apporter les amendements qui s'imposent après avoir préalablement fourni les preuves et les justificatifs qui légitiment de telles modifications.

Le processus de mesure

⁴ Perrenoud, P. (2003). Vie pédagogique

Le processus de mesure nécessite pour qu'il soit maîtrisé d'être doté d'un certain nombre de caractéristiques dont :

- l'utilisation d'instruments de mesure possédant une certaine fidélité, une certaine sensibilité, une certaine exactitude et une certaine justesse ;
- la définition correcte du principe de mesure, de la méthode de mesure, et du mode opératoire;
- la garantie de taux d'erreurs minimaux pour la répétabilité et la reproductibilité.

Cette dernière caractéristique peut être considérée comme la conséquence directe des deux premières, ce qui signifie qu'il suffit de la surveiller pour trancher sur la qualité du processus de mesure. Des auteurs comme ZAÏDI⁵ affirment que "*l'erreur globale (de mesure) peut être déterminée en combinant de façon statistique les estimations de ces deux types d'erreurs*". Il suggère par ailleurs d'utiliser la notion de capabilité pour quantifier cette erreur, et propose des méthodes pour y arriver. La capabilité étant pour lui "*une mesure non seulement de l'erreur de l'instrument de mesure lui-même mais aussi du mode opératoire, de l'opérateur et de l'environnement*".

Entre mesure et évaluation

Malgré toutes les différences qui peuvent exister entre mesure et évaluation ce qui a été dit pour la mesure reste tout à fait valable pour l'évaluation pour la raison toute simple de la relation de cause à effet qui existe entre les deux telle qu'il a été rapporté par DELANDSHEERE de la part de H.TABA qui dit : "*le processus de mesure est fondamentalement descriptif, car il indique quantitativement à quel degré un trait est possédé. La mesure en éducation est généralement concentrée sur des caractéristiques spécifiques, étroites et bien définies. L'évaluation dépend de la mesure, mais elle porte sur un profil plus large de caractéristiques et de performances*". C'est suffisamment convaincant pour s'inspirer des théories, principes et méthodes de mesure dans toute réflexion sur l'évaluation. D'autant plus que la docimologie - "*cette science qui, selon DELANDSHEERE⁶, a pour objet l'étude systématique des examens et en particulier des systèmes de notation, et du comportement des examinateurs et des examinés*" - "n'a qu'une faible influence sur les pratiques de l'évaluation" pour reprendre les propres termes de PELPEL⁷, malgré tous les efforts déployés par les spécialistes de cette science pendant plus de 60 ans. Mais, si la docimologie a relativement échoué dans l'amélioration des pratiques de l'évaluation, elle a eu le grand mérite de mettre en exergue les aspects aléatoires du processus de notation. Les travaux réalisés dans le cadre de la docimologie ont réussi à en déceler quelques unes ; on peut citer à titre d'exemple les expérimentations faites par G. NOIZET et J.P. CAVERNI⁸ qui ont révélées l'existence de deux lacunes :

- on ne se base pas sur les mêmes critères pour évaluer ;
- on n'applique pas une échelle commune des priorités pour les critères utilisés.

L'acte d'évaluation est-il modélisable ?

Cependant, il reste de toute évidence un long chemin à faire avant de pouvoir, sinon éliminer, du moins réduire la part de l'arbitraire dans la pratique de l'évaluation. Le processus de mesure étant constitué d'éléments rationnels et d'autres qui le sont moins (le facteur humain par exemple), il est très difficile de soutenir l'ambition de l'éradication totale de toute erreur d'évaluation. Mais toujours est-il que même l'erreur imputée au facteur humain, il n'est pas impossible de la rendre insignifiante pour peu que le comportement qu'incarne le professeur lors de l'évaluation soit d'abord connu puis ensuite maîtrisé. Il est temps pour éviter ce problème on ne peut plus grave, de mettre en évidence tout le cheminement intellectuel entrepris par l'enseignant lors de l'évaluation, de le modéliser, de le valider et de s'en inspirer pour formaliser les règles et les procédures à appliquer dans ce domaine.

⁵ ZAÏDI, A. 1992. SPC : Concepts, méthodologies et outils. Paris : TEC & DOC – LAVOISIER.

⁶ DELANDSHEERE, G. 1980. Evaluation continue et examens, précis de docimologie. Paris : Fernand - Nathan.

⁷ PELPEL, P. 1986. Se former pour enseigner. Paris : BORDAS.

⁸ NOIZET, G. et CAVERNI, J.P. 1978. Psychologie de l'évaluation scolaire. P.U.F.

Il est souhaitable de vérifier la validité de chaque instrument d'évaluation auprès d'un échantillon d'élèves, mais si cette opération s'avère irréalisable il est recommandé de suivre par une carte de contrôle l'évolution de sa capacité et de sa justesse afin de pouvoir remédier à temps aux problèmes qui risquent de survenir. Chaque problème nécessite la réunion de la cellule qualité pour en étudier les causes, chercher les moyens de les éliminer et rédiger les procédures à suivre pour éviter qu'ils se reproduisent à l'avenir.

Mise en œuvre

La conception des instruments d'évaluation doit être régie par des règles élaborées par l'équipe pédagogique. Un plan générique d'évaluation sera établi pour chaque entité du système éducatif (élève, enseignant, manuel scolaire, lycée, etc.) puis il sera décomposé en plans spécifiques pouvant s'accommoder aux besoins spécifiques des différents éléments d'une entité. L'élaboration du plan générique est du ressort des parties prenantes de chaque entité, par contre chaque plan spécifique est établi par les parties concernées par l'élément à qui il est destiné. Ainsi par exemple peut-on élaborer une grille d'évaluation générique pour l'évaluation de tous les enseignants sur laquelle les inspecteurs de chaque discipline peuvent se baser pour établir leur grille d'évaluation spécifique. Cette façon de procéder a l'avantage d'homogénéiser les pratiques d'évaluation, de favoriser les échanges d'expériences, d'ancrer les pratiques de la coordination entre les différentes parties et de préserver les spécificités qui peuvent exister.

L'équipe chargée de l'élaboration du plan générique est tenue d'établir les instructions d'utilisation et de les diffuser au près des utilisateurs du plan afin de les aider dans l'application. Les informations provenant des utilisateurs des différents plans spécifiques serviront pour mettre à jour et améliorer le plan générique.

Conclusion

L'activité de l'évaluation a connu une évolution importante ces dernières années. Le triomphe du cognitivisme sur le béhaviorisme a permis de donner le rôle qu'il se doit au sujet évalué et de mettre en exergue le rôle formateur de l'évaluation. Ce n'est plus un moyen pour constater les dégâts, c'est plutôt un outil efficace d'amélioration de la qualité de l'éducation. Il reste cependant des progrès à faire au moins à deux niveaux. Au niveau des incertitudes de mesure révélées par les innombrables études docimologiques mais sans que l'on arrive à proposer des vrais remèdes pour les réduire, puis au niveau de l'intégration des différentes évaluations que connaît le système éducatif en une évaluation systémique capable d'éviter les redondances et de capitaliser les améliorations. Cette dernière se révèle, entre autres, un instrument efficace de concertation entre les différents acteurs de l'éducation.