

Jacques AUBRET

VALEURS ET ÉVALUATION EN FORMATION

Résumé : Si évaluer c'est porter un jugement de valeur, alors l'acte d'évaluation « énonce » des échelles de valeur et structure les relations évaluateurs-évalués. En analysant les pratiques d'évaluation dans un système de formation, on se donne les moyens de reconnaître les systèmes de valeurs qui les sous-tendent et les styles de relations qu'elles engendrent. Notre objectif est d'attirer l'attention sur les logiques sous-jacentes aux pratiques d'évaluation des systèmes éducatifs, de l'efficacité des enseignants et des acquis des élèves. L'exposé est organisé autour de la notion de validité : à chaque outil ou procédure correspondent des jugements de valeur qui leur donnent une pertinence et une utilité. Dans cette perspective nous distinguons trois logiques de validation : validation scientifique, validation sociale, validation écologique.

Mots-clés : Évaluation psychopédagogique - Validité scientifique - Validité sociale - Bilan de compétences.

On ne peut s'interroger sur les valeurs en éducation et en formation sans réfléchir sur les pratiques d'évaluation par lesquelles s'expriment les échelles qui énoncent les valeurs dans le système éducatif, car « évaluer » c'est, selon une définition courante, porter un jugement de valeur. Expliciter les pratiques d'évaluation c'est donc se donner les moyens d'explicitier les valeurs en acte dans nos pratiques d'éducation. C'est l'objectif que je me suis donné en proposant une contribution à ce colloque.

Je ne présenterai donc pas des résultats de recherches. Néanmoins, les différents aspects de l'évaluation que je soulèverai correspondent à un ensemble de travaux personnels ou collectifs engagés dans les années 1980, dans la tradition de la psychométrie française (élaboration de tests psychologiques à l'usage des psychologues de l'éducation et de tests de sélection) et qui m'ont conduit aujourd'hui à m'intéresser aux pratiques de bilans de compétences et aux procédures de reconnaissance et de validation des acquis.

L'évaluation apparaît d'abord sous sa diversité : la « notation » dans les évaluations des productions scolaires, de la maternelle au supérieur (notation mémorisée dans les livrets scolaires) et les examens, les « appréciations » dans les bulletins scolaires (fruit des échanges — ou des rapports de pouvoir — en conseil de classe), les « épreuves standardisées » dans les enquêtes ministérielles ou les tests proposés par les psychologues et conseillers d'orientation, l'« autoévaluation » dans certaines pratiques d'aide à la prise conscience des procédures d'acquisition ou de résolution

des problèmes qui accompagnent les apprentissages, la « reconnaissance de la valeur de preuves d'acquis » dans les procédures de validation d'acquis, etc. Diversité d'objectifs également, parmi lesquels la régulation des apprentissages quand il s'agit des pratiques quotidiennes d'évaluation en cours de formation, l'orientation et la sélection scolaire quand nous consignons les notes dans un livret scolaire ou lorsque nous prononçons des résultats d'examen, l'élaboration de diagnostics ou de pronostics de réussite pour les psychologues, le pilotage institutionnel des systèmes de formation quand il s'agit d'enquêtes.

Notre interrogation voudrait dépasser ces observations de surface pour questionner les logiques d'attribution de valeur qui sous-tendent les objectifs d'évaluation que l'on se donne et les pratiques qui les mettent en musique en quelque sorte. De ce point de vue nous distinguerons trois systèmes de référence invoqués assez communément pour justifier de la validité des pratiques d'évaluation : la référence scientifique, la référence sociale, la référence écologique.

LA RÉFÉRENCE SCIENTIFIQUE COMME GARANTIE DE FIABILITÉ DES OUTILS ET PROCÉDURES

Il existe une tradition de recherches en matière d'évaluation qui indique bien que les processus et les produits de l'évaluation peuvent être éprouvés par le scientifique. La psychométrie depuis le début du siècle, la docimologie avec H. Piéron, l'édu-métrie avec J. Cardinet et les travaux de l'ADMEE (Association pour le Développement des Méthodologies d'Évaluation en Éducation) créée il y a une dizaine d'années illustrent, entre autres, cette affirmation.

Dans cette perspective, l'élaboration et le contrôle des pratiques d'évaluation obéissent aux habitudes de fonctionnement de la pensée scientifique. Chaque outil et procédure est éprouvée lors d'expérimentations comme on teste des hypothèses. Des ouvrages récents explicitent la méthodologie de la validation (Dikes et al. 1994 ; Laveault, Grégoire, 1997). Les résultats des expérimentations sont publics et discutés par la communauté scientifique.

La rigueur d'une méthode scientifique devrait contribuer à la réduction des facteurs qui affaiblissent la recherche de l'équité entre les élèves ou les candidats devant l'évaluation. Noizet et Caverni (1978), après H. Piéron (1963), démontent les mécanismes qui favorisent les biais et les dysfonctionnements de l'évaluation dans la correction des copies. Les enquêtes conduites sur les systèmes éducatifs dans le monde par différentes instances nationales ou internationales recherchent et affichent des garanties de cette nature (voir à ce sujet la présentation qui en est faite par G. de Landsheere, 1994). Les questionnaires et les épreuves standardisées destinés à collecter les informations sur les acquis scolaires sont soumis à des études de validité préalablement à toute utilisation. On ne voit pas très bien quel sens pourrait être donné à des comparaisons de résultats à tous les niveaux (classes, établissements, pays, etc.) sans cette rigueur méthodologique. La démarche scientifique permet ainsi de construire des « objets » susceptibles d'être perçus de la même manière par tous ceux qui acceptent les présupposés de cette construction.

Ces présupposés sont identifiables. Le principe de « la mesurabilité », comme possibilité d'application des propriétés des nombres à des objets qui n'ont pas nécessairement ces propriétés, est un élément clé de l'opérationnalisation de la démarche. Les études de validation permettent de ce point de vue de contrôler en partie les effets de l'intrusion de la mesure, et donc de guider les interprétations.

Les données produites par les enquêtes nationales et internationales constituent des points de repères utilisables pour situer des performances individuelles ou collectives dans des ensembles plus vastes, ou pour comparer des performances d'individus ou de groupes. On peut ainsi apprécier les effets différenciés sur les performances des élèves de traitements pédagogiques variés, des politiques éducatives, de caractéristiques individuelles ou collectives, fixer des objectifs, contrôler les programmes de formation. Ces données conditionnent dans une large mesure les regards que nous portons sur les systèmes éducatifs avec un risque évident, celui de transformer en objectif de formation et d'éducation la volonté de réduire les différences ou les écarts constatés. La différence mesurée tient lieu de valeur.

La recherche de garanties scientifiques a ses limites et ses effets pervers. Les limites sont celles que fixe la possibilité d'objectivation et de mesure des phénomènes à observer. De ce point de vue, l'évaluation des acquis se cantonne le plus souvent dans l'observation de résultats ou de performances. Les effets pervers proviennent essentiellement des extensions incontrôlées des usages des données recueillies. S'il faut faire confiance aux chercheurs qui élaborent des outils sophistiqués, il est rare que l'utilisateur puisse s'approprier la démarche dans ses aspects les plus techniques, ce qui est pourtant indispensable pour comprendre les significations des résultats produits. On risque ainsi de perdre le sens de la relativité et du caractère construit et réducteur de la démarche et des résultats.

Si nous prenons le cas de l'illettrisme en France, des mesures répétées, depuis 1983 notamment, aboutissent à des conclusions convergentes : un Français sur cinq serait incapable de se servir de la langue pour s'exprimer et comprendre (cf. Aubret, 1989). Mais les bases sur lesquelles ces informations sont construites varient d'une étude à une autre. Toutefois, ce qui semble conforter la solidité de la mesure (le fait qu'elle soit répétable sur des occasions variées) perd son sens dès que l'on veut traduire ces résultats en termes conceptuels (qu'est-ce qu'un illettré ?) ou en termes d'objectifs de remédiation (que peut-on faire pour réduire le phénomène de l'illettrisme ?), ou en termes de processus (comment conduire une personne donnée vers une lecture efficace ?).

Les réponses à ces questions impliquent la recherche d'autres systèmes de référence, et l'adhésion des acteurs concernés à leur valeur informative et communicative. Si la mesure n'épuise pas la recherche scientifique (les chercheurs élaborent aussi des modèles qui permettent de donner sens à d'autres observations que les performances) la mesure est confortée par l'importance qu'elle joue dans les représentations sociales des acquis scolaires. L'évaluation ne peut être regardée seulement sous ses aspects techniques et procéduraux. Elle se construit socialement.

En bref, la référence à la « validité scientifique » des outils et des pratiques est référence à des formes de pensée et à des méthodes de résolution de problèmes

éprouvées au sein de la communauté scientifique. La science construit des objets d'évaluation, détermine les méthodes d'observation et de mesure et fournit les cadres d'interprétation des observations effectuées. L'intérêt de telles observations réside dans leur extériorité par rapport à la personne observée et dans la non implication recherchée des observateurs. Cependant l'observation se réduit à ce qui est objectivable et les résultats obtenus dans de telles conditions sont davantage considérés comme arguments dans des jugements de probabilité d'adaptation à des situations futures que comme preuves d'acquis ou de compétences.

L'ANCRAGE SOCIAL DES PRATIQUES COMME SOURCE DE VALEURS

En matière d'évaluation, on se réfère parfois à la notion de validité sociale. Cette notion reste encore assez floue. Néanmoins nous savons tous que les échelles de valeur qui fondent nos jugements de valeur sont généralement partagées par d'autres que nous. La validité sociale exprime une certaine forme de consensus, dans des groupes donnés (en fonction des différentes formes d'appartenance des individus), sur l'attribution de valeurs à des objets d'évaluation. L'étendue des groupes n'est pas définie, pas davantage que ne le sont les critères d'attribution des valeurs. Néanmoins on constate que des pratiques, par exemple des pratiques de recrutement, sont reconnues comme valides à partir de preuves de validité faisant appel à des intuitions, des convictions, des opinions, des croyances, voire des évidences difficiles à remettre en question. On affirme sans démontrer le plus souvent en raison de la position sociale dans laquelle on est. Les mécanismes sociaux qui conduisent à ce consensus s'appuient essentiellement sur les représentations sociales partagées et les rapports de force entre individus.

On connaît par exemple, dans le domaine de la formation, l'importance de la notation. En fait, nous devrions nous interroger sur les représentations qui y sont associées et les comportements individuels et sociaux auxquels la notation donne lieu. Tout se passe comme si la note représentait une valeur vraie de l'individu, une valeur absolue. La notion d'absolu l'emporte souvent sur les considérations de relativité dans les discussions sur la valeur de référence de la moyenne, du zéro éliminatoire, des notes maximum rarement attribuées. Comment expliquer autrement que le système de la « notation » largement pris en défaut par les travaux scientifiques persiste dans les pratiques ? Des travaux sont en cours, sur le fonctionnement des évaluateurs, sur l'élaboration de ces normes internes qui guident les jugements des évaluateurs. On notera de ce point de vue les approches développées dans le cadre de la psychologie sociale, approches assez bien représentées dans les communications d'un récent colloque sur les compétences (Brangier *et al.*, 1997).

On le voit le social prend autant d'importance que le scientifique pour rendre compte de la persistance des pratiques. La notation n'exprime pas la valeur des personnes ; elle est le moyen de signifier dans le processus d'évaluation les échelles de valeur élaborées socialement. La notation assure la reproduction des modes de savoir et des formes de transmission des savoirs. Elle contribue à maintenir les distan-

ces (les positions respectives) entre les évaluateurs et les évalués. Une question d'actualité : sur quelles bases les étudiants de l'université évalueront-ils leurs enseignants ?

Il n'y a pas nécessairement opposition entre validité sociale et validité scientifique (d'une certaine façon l'expert scientifique partage un certain pouvoir) mais l'opacité et la complexité du recours au rationnel scientifique (du fait des méthodes ou des objets évalués) peut faire le lit d'une validation sociale sauvage. La recherche de validité sociale ne peut être une simple alternative à la recherche d'un rationnel scientifique comme garantie de la fiabilité des pratiques d'évaluation. Rationnel scientifique et validité sociale peuvent représenter deux facettes interactives d'une même exigence. On peut espérer que l'acceptabilité sociale sera d'autant mieux assurée qu'elle pourra s'appuyer sur des indicateurs d'efficacité du sujet et que les preuves objectives seront d'autant mieux reconnues qu'elles prendront en compte l'intelligibilité sociale (l'appropriation par les acteurs concernés) des méthodes par lesquelles elles sont produites.

LA RECONNAISSANCE DES SIGNIFICATIONS ET DES VALEURS PRODUITES PAR LE SUJET

S'intéresser au sujet comme source de jugements de valeurs sur des objets qu'il est le seul à appréhender c'est retourner sur les lieux mêmes de réalisation des processus de formation et d'éducation. La notion de validité écologique (que le monde des chercheurs s'approprie, voir H. Gardner 1996, par ex.) exprime cette idée et la nécessité de réhabiliter le sujet (la personne) dans le processus même d'évaluation. On met notamment l'accent sur l'importance du travail de « reconnaissance » de soi. Ainsi l'évaluateur ne s'intéresse pas seulement aux occasions de formation ou d'expérience, aux performances observables, mais il prend en compte les significations que leur accorde le sujet lorsqu'il traite ces expériences en vue de les constituer en savoirs. Le sujet est donc partie prenante de son évaluation. Auto-formation et autoévaluation constituent une même réalité (voir à ce sujet les développements le développement du thème dans la formation des adultes dans Carré et Pearn, 1992).

Le champ de la validation des acquis des expériences sociales et professionnelles, et le domaine couvert par les pratiques de bilan de compétences illustrent la nécessité d'une telle approche de l'évaluation. L'expérience professionnelle nous renvoie vers le sujet et son activité cognitive organisatrice. Les acquis de l'expérience ne se situent pas dans le prolongement des apprentissages organisés (et validés par des diplômes) mais comme le résultat de ce que la personne découvre, apprend, organise par elle-même dans les occasions de vie et de travail qu'elle crée ou rencontre. Comment peut-on les reconnaître en dehors d'une participation active de la personne ?

La reconnaissance du sujet comme source de valeur constitue l'un des aspects majeurs de la définition des pratiques de bilan de compétences : « les actions permettant de réaliser un bilan de compétences... ont pour objet de permettre à des

travailleurs d'analyser leurs compétences professionnelles et personnelles ainsi que leurs aptitudes et leurs motivations afin de définir un projet professionnel et, le cas échéant, un projet de formation » (loi du 31 décembre 1991). Tout se passe comme si l'on acceptait l'idée qu'il faut observer, analyser, étudier, comprendre les processus d'acquisition sur les lieux mêmes où ils sont à l'œuvre et avec les acteurs directement concernés. La prise en compte de la « validité écologique » est reconnaissance que la personne ne peut être dépossédée de son pouvoir de donner du sens et de la valeur.

QUELLES ÉVALUATIONS POUR QUELLES VALEURS ?

Il se peut qu'il y ait des décalages entre les discours sur les finalités de l'éducation, notamment lorsque l'on s'exprime en terme de valeurs (les valeurs républicaines, les valeurs citoyennes, etc.) et ce que véhiculent les pratiques en termes d'échelles de valeur. Dans ses pratiques, l'école n'est pas neutre. L'analyse des pratiques d'évaluation montre l'évaluation comme un lieu de tension où s'affrontent idéologies et acteurs. Les pratiques d'évaluation sont plurielles et pluridimensionnelles.

Si la référence scientifique est considérée comme pertinente c'est parce qu'elle protège l'homme de sa subjectivité. On pense que l'évaluation scientifique devrait contribuer à accroître l'équité entre les individus évalués. Il y a cependant du chemin à parcourir. Mais ce qui pourrait devenir une religion de l'objectivité pourrait paradoxalement construire un homme dépouillé de sa véritable nature : un « sujet » responsable de ses actes (y compris cognitifs) dans une société de semblables. Les considérations sur la notion de validité sociale, que l'on devrait approfondir, introduisent dans le débat sur l'évaluation la référence à l'intersubjectivité. L'affrontement des valeurs dans la collectivité et les représentations qu'elles engendrent chez les sujets imprègnent les pratiques sociales. Dans une société qui mise sur l'éducation pour faire émerger des sujets responsables et autonomes le développement des capacités d'analyse des interactions sociales pourrait être une voie éducative : les échelles de valeur ne seraient pas dictées mais construites contre des valeurs imposées.

La pluralité des sources de valeurs ancrées dans les pratiques est un fait. Loin de le déplorer, nous pensons que chacun des aspects évoqués ci-dessus (validité scientifique, validité sociale, validité écologique) évoque des sources d'attribution de valeurs que l'on devrait pouvoir faire interagir. Si la visée éducative est d'instaurer l'homme comme « valeur » dans une communauté humaine qu'il contribue à construire en se réalisant, la subjectivité devient un obstacle lorsqu'elle rend la communication interhumaine impossible. De la même manière, l'objectivité devient un obstacle lorsqu'elle construit des observables qui masquent l'auteur et les intentions qui les ont produits. Enfin la collectivité perd sa raison d'être dans cette visée éducative quand elle cesse d'être l'arbitre et le médiateur entre ces deux pôles. Ce qu'elle fait quand les pratiques d'évaluation, au nom de la mesure objective discri-

minent, disqualifient et excluent, quand le système de valeurs d'un individu (enseignant, chercheur, chef d'entreprise, ministre,...) ou d'un groupe devient une référence absolue, ou quand elle promeut des pratiques qui échappent à la fois au contrôle scientifique et aux décisions humaines (pratiques ésotériques notamment).

Jacques AUBRET

Université Charles de Gaulle — Lille 3

Summary : If to evaluate is to make a value judgement, the act of evaluation « evinces » scales of value and structure of evaluators-evaluated relations. Analysis of evaluation practices in a training system provides the means to recognise the systems of values on which they are based and the style relations they engender. Our objective is to draw attention to the approaches underlying evaluation practices in educational systems, evaluation of effectiveness of teachers and proficiency acquired by students. The commentary is arranged around the concept of validity : each tool or procedure is concerned by value judgements which give them pertinence and usefulness. This leads us to identify three approaches to validation : scientific validation, social validation, ecological validation.

Keywords : Psychoeducational evaluation - Scientific validity - Social validity - Ecological validity - Skills assessment.

Bibliographie

- Aubret J. (1989) « Être ou devenir illettré ? » *Humanisme et Entreprise*, Cahiers du Centre d'Études et de Recherches 173 (1-13).
- Aubret J. & Gilbert P. (1994) *Reconnaissance et validation des acquis*. Paris : PUF.
- de Landsheere G. (1994) *Le pilotage des systèmes d'éducation*. Bruxelles : De Boeck.
- Brangier E., Dubois N. & Tarquino C. (eds) (1997) *Compétences, contextes professionnels*. Colloque international, Metz, 19-20 juin 1997.
- Cardinet J. (1987) *Évaluation scolaire et mesure*. Bruxelles : De Boeck.
- Carré P. & Pearn M. (1992) *L'autoformation dans l'entreprise*. Paris : éditions Entente.
- Dickes P., Tournois J., Flieller A. & Kop J.-L. (1994) *La psychométrie*, Paris : PUF.
- INED-INETOP (1978) « Enquête nationale sur le niveau intellectuel des enfants d'âge scolaire. » *Travaux et Documents*, cahier 83.
- Gardner H. (1996) *Les intelligences multiples*. Paris : Retz. (trd. fr. de *Multiple Intelligence. The Theory in Practice. A reader*. Harper Collins Publishers).
- Hadji C. (1992) *L'évaluation des actions éducatives*. Paris : PUF.

J. AUBRET

Laveault D. & Grégoire J. (1997) *Introduction aux théories des tests en sciences humaines*, Bruxelles : De Boeck.

Noizet G. & Caverni J.-P. (1978) *Psychologie de l'évaluation scolaire*. Paris : PUF.

Piéron H. (1963) *Examens et docimologie*. Paris : PUF.