

- le cours théorique et ses exemples (pages 1 à 16)
- les références bibliographiques (pages 17 à 20)
- le glossaire des concepts-clés (pages 21 à 23)

## QUELQUES ELEMENTS FONDAMENTAUX SUR L'EVALUATION

Il s'agit ici de **l'évaluation instituée** qui s'appuie sur une instrumentation spécifique.

J.M. Barbier (1985) distingue l'évaluation instituée de l'évaluation implicite, qui ne se révèle qu'à travers ses effets (décider de reprendre une explication, décider de ne pas se présenter à un examen), et de l'évaluation spontanée, qui ne se traduit que par la formulation d'un jugement intuitif (c'est souvent le cas des jugements « de couloir » : un tel est très bon ou pas bon...).

**L'évaluation sert à produire de l'information « éclairante » pour prendre des décisions.** Les modalités de l'évaluation sont fonction du type de décisions à prendre.

Après avoir défini en quoi consiste l'évaluation, nous en distinguerons les différentes fonctions, et enfin nous présenterons quelques principes et perspectives pour mettre en œuvre une évaluation au service des apprentissages.

### 1. QU'EST-CE QU'EVALUER ?

Les réponses à cette question ont donné lieu à **une abondante littérature**. Théories et modélisations de pratiques ont modifié la façon de penser l'évaluation scolaire. **Les instructions officielles ont pris en compte en partie les résultats de ces travaux en insistant sur la dimension formative de l'évaluation.** Mais les pratiques installées ne relèvent pas toujours d'une démarche rigoureuse de l'évaluation, et les « règles du jeu » n'y sont pas toujours clairement explicitées.

## 1.1. Mettre en relation un référé et un référent pour se prononcer sur...

Toute évaluation suppose une mise en relation d'**un référent** et d'éléments, issus d'un observable -**un référé**- (Barbier, 1985), pour **produire de l'information éclairante** sur l'observable (Hadji, 1990). L'évaluation est **un processus** qui suppose l'élaboration d'un référent et la détermination du référé. L'évaluation peut porter sur un produit (une réalisation de l'élève) et/ou un processus (la démarche utilisée).

### 1.1.1 Evaluer une production

L'évaluation scolaire met plutôt l'accent sur l'évaluation de productions réalisées par les élèves, en fonction de consignes, en vue de se prononcer sur le développement de compétences visées.

**Le référent** comprend alors l'identification de **la/les compétence/s visée/s et les qualités ou critères de réussite** du produit attendu, résultant d'une tâche prescrite, et témoignant de la mise en œuvre de cette / ces compétence/s.

Exemples : les compétences, comme « résumer un texte, rendre compte de , argumenter, analyser et commenter des documents... » seront évaluées à travers des tâches écrites ou orales données à réaliser aux élèves. Types de tâches : résumés d'un texte narratif/informatif/argumentatif, comptes rendus d'une manifestation/d'une réunion..., débat oral, discussion écrite à partir de documents divergents, dissertation, commentaires de documents variés portant sur le même thème, commentaire composé... Le choix des tâches, des supports et des degrés d'exigence sont fonction des compétences référencées pour le niveau des élèves visés et de la période d'apprentissage.

Les critères qui définissent l'attendu ou critères de réussite sont fonction du type de tâche prescrite. Ils ressortent de quelques grandes catégories :

- conformité aux consignes, exactitude des résultats / des calculs, des informations utilisées ... / complétude des informations fournies, des réponses apportées... ;
- cohérence d'une argumentation, d'un raisonnement... ;
- pertinence d'une explication, des idées développées, des exemples fournis, des choix d'arguments / d'informations / de modèles utilisés, de la problématique, de l'organisation... ;
- communicabilité de l'expression -correction de la syntaxe de l'orthographe, concision de l'expression, lisibilité de l'écriture et de la présentation- ;
- originalité des idées, de l'expression, des exemples...

Exemple de critères de réussite pour un résumé :

- concision (qui correspond à la catégorie conformité, autrement dit : le texte produit correspond à la longueur indiquée, il s'agit bien d'un résumé)
- fidélité (ou pertinence du contenu, c'est à dire qu'on retrouve les idées essentielles et la structure de base)
- cohérence
- communicabilité ou correction de la syntaxe et de l'orthographe et lisibilité de l'ensemble
- reformulation personnelle

Les critères de la production attendue sont choisis et hiérarchisés en fonction des compétences visées et du moment de l'évaluation dans la séquence d'enseignement-apprentissage.

**Le référé**, c'est **ce qui est observé** (la production réalisée), **saisi à travers un faisceau d'indicateurs** qui permettent de se prononcer en fonction des critères. Les indicateurs sont la version ou conversion concrète d'un critère, un indice pris dans l'observable qui permet de dire si l'objet répond au critère. Un indicateur peut correspondre à plusieurs critères, plusieurs indicateurs, se combinant, peuvent correspondre à un même critère.

Exemples d'indicateurs pour le résumé:

- la longueur du résumé produit est significative du respect de la consigne, de la concision, mais pas exclusivement ;
- le nombre d'informations / d'idées ou d'événements apparaissant dans le résumé est significatif de la complétude, à condition qu'ils correspondent à ce qui se trouve dans le texte de base ;
- l'enchaînement des informations / des idées ou événements, les connecteurs utilisés sont significatifs de la cohérence...

Il faut le plus souvent **croiser plusieurs indicateurs** pour se prononcer sur une dimension de la production, comme on croisera plusieurs critères pour se prononcer sur la qualité globale de la production. Par exemple, pour juger de la fidélité d'un résumé par rapport au texte de base, on regardera si le nombre d'idées ou d'événements repris dans le résumé correspond au nombre d'idées du texte à résumer, si la structure choisie pour le résumé et les marques d'énonciation utilisées sont les mêmes que celles du texte de base... Pour juger de sa pertinence, on regardera si les informations / idées / événements repris sont ceux qui sont essentiels dans le texte de base, si les marques d'énonciation correspondent à la fonction du résumé (synthèse ou présentation d'un texte informatif / argumentatif / narratif).

Construire un référent pertinent et définir des indicateurs correspondant aux critères est gage d'une évaluation rigoureuse.

### *1.1.2. Evaluer une démarche*

Pour **évaluer une démarche**, l'**accent** n'est pas mis sur les critères de réussite mais **sur les critères de réalisation**, c'est à dire les opérations invariantes impliquées par la réalisation de la tâche<sup>1</sup>, qui ressortent d'une analyse de la tâche.

Exemple pour le résumé :

- Identifier le type de texte à résumer, la fonction du résumé et le destinataire
- Distinguer idées et illustrations de ces idées,
- Dégager la structure du texte,
- Sélectionner les informations / idées événements essentiels,
- Les reformuler,
- Organiser le résumé en reprenant la structure et (suivant le fonction du résumé) les marques d'énonciation,
- Rédiger l'ensemble de façon personnelle et concise.

L'ensemble de ces critères de réalisation ne se confond pas avec une méthode (dont les opérations sont ordonnées). Chaque réalisateur organise et met en œuvre les opérations en fonction de ses propres processus cognitifs.

### *1.1.3 Les 3 intentions de l'évaluation et la communication des résultats*

Hadji (1990, 1992) distingue l'intention d'**estimer**, d'**apprécier**, de **comprendre**. Quelle que soit l'intention, une évaluation rigoureuse nécessite des critères. Mais avec l'intention d'**estimer et/ou d'apprécier**, l'accent sera mis sur la production de l'élève et on privilégiera **les critères de réussite**. Avec l'intention de **comprendre**, l'accent sera mis sur la démarche mise en œuvre par l'élève et on privilégiera alors **les critères de réalisation**.

La communication des résultats de l'évaluation prend des formes différentes, utilise des supports et codes différents, s'adresse à des destinataires différents suivant l'intention dominante et les partenaires du « jeu » (enseignant-élève/élèves ou enseignant-élève-parents-institution).

L'évaluation qui a pour **intention de comprendre**, si elle peut se traduire par des appréciations (difficilement par des estimations car elles signifient une fin au moins provisoire d'apprentissage et donc n'incitent pas à s'y remettre), va plutôt s'exprimer sous forme de

---

<sup>1</sup> On trouvera des exemples de critères de réalisation dans les livres de Veslin (1992) et Grandguillot (1993), dans le livre coordonné par M. Grangeat (1997), p.46.

**questions relatives à la mise en œuvre des critères de réalisation, à la représentation que se fait l'élève de la tâche et de son but.** Elle concerne l'enseignant et l'élève/les élèves.

L'évaluation estimative et/ou appréciative d'une production, d'un niveau de compétences relève du **jugement d'expert**, qui peut un temps ne s'exprimer qu'entre les partenaires de la classe, mais qui est destinée à être communiquée aux autres partenaires : les parents et l'institution. Les modes de communication de ce jugement inscrit sur la production elle-même, sur un bulletin d'évaluation, varient. **Traditionnellement une note chiffrée traduit le quantitatif, des mots (appréciations, annotations) expriment le qualitatif.** Mais sont aussi utilisés des symboles (figure qui sourit ou grimace, couleurs...), des + ou des - dans une grille de critères ou de compétences ou les codes A (acquis) / EC (en cours d'acquisition) / NA (non acquis).

La note chiffrée, utilisée officiellement dans le secondaire (Cf. les nouveaux bulletins des collèges) et plus officieusement dans des écoles primaires, invite à la comparaison entre élèves et au classement. Elle relève davantage d'une évaluation normative que d'une évaluation critériée.

**Plus novatrice est la pratique du port-folio**, qui rassemble des travaux choisis par les élèves pour témoigner de leurs apprentissages et est le support d'un entretien élève-enseignant-parents<sup>2</sup>.

## **1.2 Tendances normative et « parasitages » de la notation chiffrée « traditionnelle »**

L'usage de la notation chiffrée pour traduire le résultat de l'évaluation et le flou des pratiques courantes d'évaluation dans les classes (souvent sous prétexte de gagner du temps) amènent un usage non pertinent de l'évaluation normative.

On appelle évaluation normative une évaluation « dont le cadre de référence est constitué par les scores obtenus, lors d'une épreuve semblable, par les membres d'un groupe de référence » (Hadji, 1990). C'est le type d'évaluation utilisée en psychologie par exemple pour les tests de mesure du QI. **« Par extension, est dite normative une évaluation dont l'ambition est de situer les individus les uns par rapport aux autres.** Evaluation normative s'oppose alors à évaluation critériée : ... évaluation qui apprécie un comportement en le situant par rapport à une

cible [le critère qui correspond à l'objectif à atteindre] » (Hadji, 1997, p. 15). On peut ainsi qualifier de normative l'évaluation qui débouche sur une notation chiffrée.

En effet, la note ressemble à une mesure sans en être une, car « une mesure est objective, en ce sens que, une fois définie l'unité, on doit toujours avoir la même mesure du même phénomène » (Hadji, 1997, p. 24), quel que soit le mesureur. Or, comme nous allons le voir, **la note varie en fonction de l'évaluateur**. De plus, la mesure ne prend du sens que par comparaison avec un « étalon », d'autres mesures. Or, l'enseignant, quand il évalue la qualité d'une production, un niveau de compétences, ne dispose pas d'un « étalon ». Aussi, à défaut, il compare les productions<sup>3</sup>, pour ventiler les notes sur l'échelle traditionnellement adoptée de 0 à 20<sup>4</sup>, le plus souvent suivant une courbe de Gauss. Ainsi, **la note attribuée est fonction de l'interprétation que fait l'enseignant de l'écart qu'il constate entre la production et un référent plus ou moins explicite<sup>5</sup>**, qui peut varier en fonction du contexte de l'évaluation.

Les docimologues<sup>6</sup> (spécialistes de l'étude de la notation) ont montré, depuis les années 50, que la notation varie en fonction des évaluateurs (Cf. Noizet & Caverni, 1978, Merle, 1998). Ils ont mis en évidence que l'estimation des performances est parasitée<sup>7</sup> par **différents effets dont l'évaluateur n'a pas forcément conscience** :

- Effets d'ordre (les premières copies sont évaluées plus sévèrement) ;
- Effets d'ancrage (surévaluation ou sous-évaluation selon que la copie se situe immédiatement après une copie jugée mauvaise ou bonne) ;
- Effets de la place des erreurs dans la copie (les erreurs en début de copie provoquent un a priori défavorable) ;

---

<sup>2</sup> Cf. le numéro thématique que lui consacre la revue *Mesure et évaluation en éducation*, Vol. 20, n° 3, 1998.

<sup>3</sup> Sur les bulletins du collège, la moyenne chiffrée des résultats de l'élève est présentée en regard de la moyenne des moyennes des élèves de la classe, de la plus basse et plus haute moyenne.

<sup>4</sup> Toutefois, la référence et le niveau d'exigences n'étant pas préalablement explicites, l'échelle a tendance à se réduire, les extrêmes étant souvent évités.

<sup>5</sup> Quand le référent n'est pas préalablement explicité, il est fait d'un attendu « à gros traits », qui tient compte des consignes, de ce que l'enseignant sait pouvoir attendre de ses élèves sur ce type de tâches, et aussi progressivement, au fil de l'évaluation de l'ensemble des productions des caractéristiques de la production moyenne, qui se dégagent.

<sup>6</sup> Docimologie, terme répandu par Piéron dans les années 60 - PIERON (H), 1963 - *Examens et docimologie*, Paris, PUF-, signifie Science et pratique du contrôle des connaissances.

<sup>7</sup> Pour réduire l'impact de ces effets parasites, on utilise dans les examens et concours diverses modalités : double correction, harmonisation des notes, pré-analyse collective d'une copie pour s'entendre sur les indicateurs ...

- Effet de source (ce que l'évaluateur sait de l'évalué ou ce qu'il en perçoit à travers la copie provoque une impression plus ou moins favorable en fonction de la subjectivité du correcteur).

Il faut ajouter que la notation chiffrée est aussi influencée, comme l'a montré Chevallard (1986) par **une négociation implicite entre l'enseignant et les élèves** (vous allez progresser en cours d'année, donc en début d'année les notes ne sont pas très élevées, ou vous n'avez pas assez travaillé ces derniers temps et je vous signifie mon mécontentement par des notes sévères...).

### 1.3 L'évaluation ne peut être ni complètement objective, ni absolue

Expliciter le référent pour évaluer les productions par rapport à lui, avoir conscience des effets parasites, permettent de diminuer, sans pour autant pouvoir l'éliminer la part de subjectivité que contient toute évaluation. **L'évaluation contient une part irréductible de subjectivité.**

De plus, la performance de l'élève est elle aussi relative à un contexte et à un moment (sinon cela signifierait que la production réalisée, l'apprentissage est terminé). On ne peut pas multiplier trop les situations d'évaluation pour une même compétence, à la même période, pour voir si la performance se stabilise. Aussi, on peut dire que **l'évaluation des compétences ne relève que d'une inférence.**

On ajoutera que si des critères sont élaborés préalablement à l'épreuve avec les élèves, ou moins leur sont communiqués, non seulement l'évaluation est plus rigoureuse, on a plus de chances d'inférer de la compétence visée que de l'aptitude de l'élève à déceler les attendus implicites de l'enseignant.

## 2. LES FONCTIONS DE L'EVALUATION

Evaluer est donc « une lecture particulière de la réalité », suivant l'expression de Hadji (1990, 1997), qui prend **différentes formes en fonction de décisions à prendre**. Le référent et le choix du référent (de ce qui est observé) dépendent de la fonction de l'évaluation qui est privilégiée (de ce à quoi va servir cette évaluation).

## 2.1 Régulation sociale et régulation pédagogique ; évaluation sommative et évaluation formative

Depuis les travaux de Scriven (1967) sur les curricula, repris par Bloom (1971) dans le contexte de l'enseignement, on considère que **l'évaluation scolaire est au service de la régulation : régulation sociale** par la gestion des flux d'élèves ou d'étudiants, **régulation pédagogique** dans le système enseignement-apprentissage. Elle assume deux types de fonctions : la fonction sommative et la fonction formative, auxquelles les chercheurs suisses ont ajouté la fonction pronostique (Cf. Allal *et al.* 1979, dans un ouvrage qui a fait date).

- ❖ « Une forme de régulation est de s'assurer que les caractéristiques des élèves correspondent aux exigences préétablies du système de formation. Dans ce cas, l'évaluation est un moyen de contrôle de la progression de l'élève aux points d'entrée, de passage et de sortie du système.
  - La fonction de l'évaluation est **pronostique** au niveau de l'accès à un cycle ou une année d'étude...
  - Sa fonction est **sommative** lorsque le contrôle s'opère à la fin d'une année d'études...
  
- ❖ Une autre forme de régulation est de s'assurer que les moyens de formation proposés par le système soient adaptés aux caractéristiques des élèves.
  - Dans ce cas, l'évaluation assume une fonction **formative** car son but est de fournir des informations permettant une adaptation de l'enseignement aux différences individuelles dans l'apprentissage » (Allal, 1979, éd. de 1985, p. 131).

L'évaluation pronostique ou sommative relèvent davantage du contrôle des apprentissages, alors que l'évaluation formative relève davantage de la régulation des conditions de l'apprentissage.

## 2.2 L'évaluation formative / formatrice ou l'approche complexe de l'évaluation

L'évaluation formative, ainsi définie dans les années 70, par ceux qui se réfèrent au **Behaviorisme** comme théorie de l'apprentissage (Bloom avec la pédagogie de la maîtrise, les tenants de la pédagogie par objectifs), vise une **régulation externe de l'apprentissage** (Cf. Bloom, 1998, Allal, 1979/1985, 1988).

Les chercheurs en évaluation, qui se réfèrent aux théories **socio-constructivistes** de l'apprentissage, complexifie l'évaluation formative en visant une **régulation interne** de l'apprentissage. Pour eux, l'élève, acteur principal de son apprentissage, doit aussi devenir l'acteur principal de l'évaluation de son apprentissage. L'évaluation prend une dimension **formatrice ou autoévaluative**, suivant la terminologie utilisée par les chercheurs d'Aix-Marseille (Bonniol *et al.* 1978 ; Bonniol, 1986, 1988 ; Nunziati 1990 ; Vial, 1987, 1995). La construction de la référence est aussi l'affaire de l'élève, d'une référence qui lui permette de se représenter la performance à réaliser mais aussi les moyens d'y arriver.

J.J. Bonniol (1986, 1988, 1989) alors distingue critères de réussite et critères de réalisation (Cf. plus haut). Les deux types de critères constituent une référence pour agir / réaliser, **une référence évolutive**, en fonction de la progression de l'apprentissage, avant d'être une référence pour l'évaluation sommative. Cette **référence partagée** entre les élèves et l'enseignant, qui évolue au fil des évaluations intermédiaires, permet à l'enseignant d'identifier réussites et erreurs des élèves, de mettre en place des régulations interactives, rétroactives et proactives (Allal, 1988, 1991) et aux élèves de se guider dans la réalisation, de réguler leur action, de mieux comprendre les appréciations de l'enseignant, ses questions.

D'une façon générale, on peut dire que **l'évaluation formative** (telle qu'elle est conçue actuellement) recouvre des stratégies qui font interagir la boucle de régulation du dispositif pédagogique, gérée par l'enseignant et la boucle de régulation de l'apprentissage gérée par l'élève. Il ne s'agit pas d'une évaluation formative qui se réduirait à du contrôle continu, qui ne viserait qu'une conformation à la référence préétablie et fixe, mais d'une évaluation qui interroge les représentations, les références implicites à l'œuvre, les démarches, qui cherche à comprendre un

fonctionnement ou un dysfonctionnement, pour en dégager le sens et orienter l'action (Hadji, 1990). C'est **une évaluation partagée qui vise l'interaction des autoévaluations des acteurs** (Campanale, 1997).

L'évaluation est pensée comme interprétation par l'évaluateur d'une réalité complexe et changeante, interprétation ponctuelle, provisoire qui cherche à éclairer un devenir possible. **L'évaluation formative englobe et dépasse le contrôle** (Cardinet, 1992). Pluri-référentielle (Ardoino, 1993), l'évaluation est essentiellement formative (Ardoino et Berger, 1986, 1989) ; elle intervient dans un système ouvert en vue d'en optimiser le fonctionnement et de dégager le sens de son évolution (Ardoino et Berger, 1986, 1989). Référent et référé se construisent, s'éprouvent et sont sans cesse dépassés et à reconstruire. L'évaluation est un processus sans achèvement (id.).

### 2.3 Six questions à se poser pour évaluer et quelques principes

Toute situation d'évaluation gagne à être éclairée par 6 questions :

- A quoi va servir cette évaluation ? Pour réguler le processus enseignement-apprentissage ? Pour contrôler des acquis ?
- A qui les résultats sont-ils destinés ? Aux élèves / à l'enseignant ? A l'institution, aux parents ?
- Qu'est-ce qui est évalué ? Des stratégies, des démarches ? Des apprentissages spécifiques / des compétences ?
- Qui évalue ? Les élèves et l'enseignant ? L'enseignant, l'équipe pédagogique / éducative ? L'institution ?
- Quand évaluer ? Avant, au début de la séquence, en cours de séquence ? A la fin de la séquence, ultérieurement en différé ? A la fin d'un cycle de scolarité ?
- Comment ? Par des observations, des analyses, des questionnements ... ? Par des tests, des épreuves ?

Le dispositif d'évaluation est à concevoir en même temps qu'est conçue la séquence enseignement-apprentissage / le projet de formation / le projet pédagogique...

**Le type de tâche pour l'évaluation sommative**, l'épreuve, qui permettra d'évaluer les compétences visées par la séquence d'enseignement-apprentissage est donc **à déterminer en même temps que sont déterminées ces compétences**. Et c'est en fonction de ces compétences et du type de tâche / d'épreuve pour l'évaluation sommative que sont organisés et le déroulement de la séquence et les évaluations intermédiaires, qui permettent de réguler le dispositif, de l'adapter en fonction de la progression des apprenants. Cela permet à l'enseignant lors du choix définitif de la tâche pour l'évaluation (adaptée à la progression dans de l'ensemble des élèves) **d'évaluer ce qui a été effectivement enseigné**.

Mais prévoir le dispositif d'évaluation ne se limite pas à envisager la tâche d'évaluation finale, c'est aussi introduire de l'évaluation formative. Il ne s'agit pas pour lui de seulement contrôler en cours de route ce qui se passe pour réguler, adapter son enseignement, renvoyer aux apprenants des informations qualitatives, mais aussi **d'impliquer les élèves dans l'évaluation** : aider les élèves, en début de séquence, à se construire une référence de ce qu'ils devront avoir appris à faire (explicitier la/les compétences visées, co-construire les critères), les mettre en situation d'utiliser les critères, d'identifier des indicateurs, de questionner leurs productions intermédiaires dans des situations d'évaluations mutuelles<sup>8</sup>, de faire évoluer les critères.

L'identification des compétences visées, la construction, l'appropriation des critères et indicateurs font partie de l'apprentissage : c'est ce qui permet à l'apprenant de réguler son action, autrement dit d'autoévaluer.

### 3. QUELQUES MOTS SUR L'AUTOEVALUATION

Tout un courant s'intéresse au développement des capacités d'autoévaluation des apprenants (Nunziati, 1990 ; de Ketele, 1986 ; Allal *et al.*, 1993 ; Paquay *et al.*, 1990, Paquay, 1994 ; Campanale, 1995, 1997) et aux stratégies visant la mise en œuvre des modalités d'autoévaluation de la part de chaque catégorie d'acteurs du système éducatif. L'évolution des

---

<sup>8</sup> Le plus souvent pour l'enseignant, les indicateurs sont si évidents qu'il n'est plus nécessaire de les formuler. Pour l'élève, nommer, reconnaître des indicateurs, en évaluant les productions d'autrui, lui permettront de comprendre les critères, d'avoir des exemples concrets de ce que signifie la formulation abstraite du critère.

théories et des modèles conduit à attribuer un **rôle déterminant à l'autoévaluation** ou à l'évaluation en première personne, suivant l'expression de Cardinet (1989) **dans la régulation et l'évolution du système enseignement-apprentissage**. L'autoévaluation a ses sources dans les travaux de Piaget (1974a, 1974b) sur la prise de conscience et dans ceux de l'école Vygotskienne (Cf. notamment Galpérine, 1980) sur le rôle du langage dans le contrôle cognitif.

### 3.1 Définitions et modélisation

L'autoévaluation a d'abord été définie comme un dialogue de soi à soi du sujet (Nunziati, 1990).

Pour Linda Allal (1993), il s'agit d'une réflexion métacognitive, qui enclenche des autorégulations : qu'est-ce que je sais, qu'est-ce que je sais faire ? Comment je m'y prends ? Qu'est-ce que je peux modifier ?

On peut actuellement la définir comme comme **un autoquestionnement** (Vial, 1997) par le sujet de son action ; **un processus d'altération du référentiel d'action d'un sujet, qui l'amène à interroger, réguler, transformer son action et par là agir sur lui** (Campanale (1997). Altération est à prendre ici au sens premier, étymologique, de rendu autre. Il exprime la nécessaire intervention d'autrui pour que le sujet puisse, aller au delà de la prise de recul et du simple constat d'écart entre le résultat de son action et l'attendu.

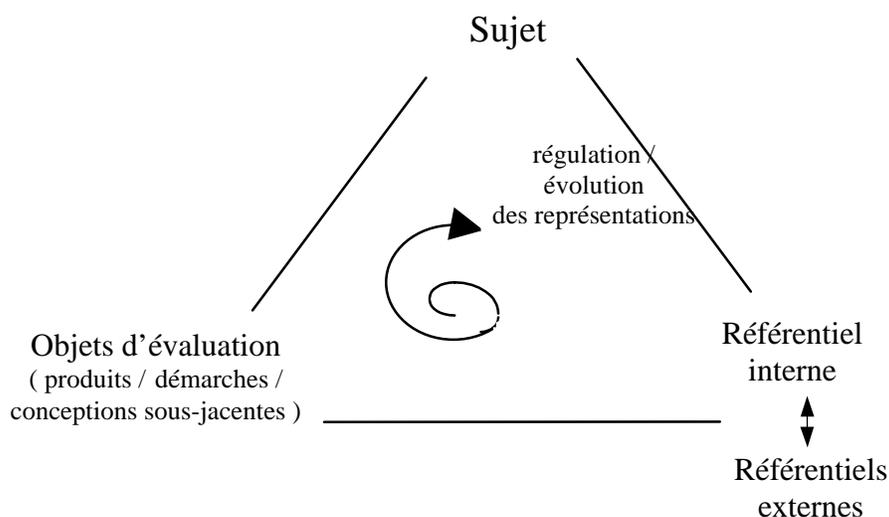


Figure 1 - Modèle de l'autoévaluation

Autrui intervient pour aider le sujet à se distancier et expliciter sa démarche, sa représentation du but, et ce faisant à conceptualiser son action. Et c'est par la **confrontation avec d'autres possibles** (produits, démarches), soumis au sujet par des pairs, avec **d'autres interprétations** des objets évalués en commun, que le sujet peut **se décentrer de son action et l'interroger**. C'est en passant par d'autres référentiels que le sujet sera amené à interroger le sien.

Ce processus spiralaire dépasse l'autocontrôle du produit de l'action pour aller à l'auto-questionnement de ses stratégies et représentations (Vial, 1997). Il débouche sur des décisions en partie conscientes et explicites de modifications de conduites en fonction de représentations modifiées ou d'arrangements avec soi-même.

### 3.2 les situations d'autoévaluation et une grille de lecture

Linda Allal (Allal et Michel,1993) distingue trois types de modalités d'organisation de situations d'autoévaluation.

- **L'autoévaluation au sens strict**, où le sujet contrôle sa production avec une grille de référence. Ce qui se passe dans la tête du sujet est difficilement observable ; un jeu se joue avec lui-même, avec les autres et avec les enjeux de la situation. Les grilles d'autoévaluation (qui fleurissent dans les manuels) peuvent avoir des effets pervers : confusion des rôles (rappelons que l'enseignant est expert), effet « caméléon » (Paquay, 1994) - l'autoévalué cherche à se conformer à l'évaluation que l'évaluateur externe pourrait faire de son travail-... Elles peuvent toutefois donner l'habitude d'autocontrôler son résultat.
- **Les évaluations mutuelles réciproques** (faites sur des productions émanant des élèves-évaluateurs) **et non réciproques** (faites sur des productions externes aux élèves-évaluateurs), qui sont celles qui permettent aux élèves de s'approprier les critères établis en commun ou par l'enseignant et déclenchent le plus de modifications.
- **Les co-évaluations** (terme plutôt réservé aux évaluations conjointes de l'élève et de l'enseignant), qui permettent d'engager un dialogue entre enseignant et élève sur les écarts d'appréciations et entrent dans une pédagogie individualisée. Elles peuvent se faire sur une production ou un ensemble de production dans une situation de bilan d'apprentissage.

La prise de distance, qu'implique l'autoévaluation, est une fonction essentielle de l'apprentissage (Berbaum, 1991, 1993), qui se répercute sur chacune des autres fonctions qu'exerce un apprenant (attitudes-projet, choix de situation, traitement des informations).

Nous avons, nous-même (Campanale, 1995, 1998, 1999) distingué **3 mouvements de prise de distance : le recul, la distanciation et la décentration**. La mise en relation des situations d'autoévaluation, des objets évalués, des mouvements de prise de distance et des régulations métacognitives qu'ils sont susceptibles d'enclencher nous a amené à proposer la grille de lecture suivante :

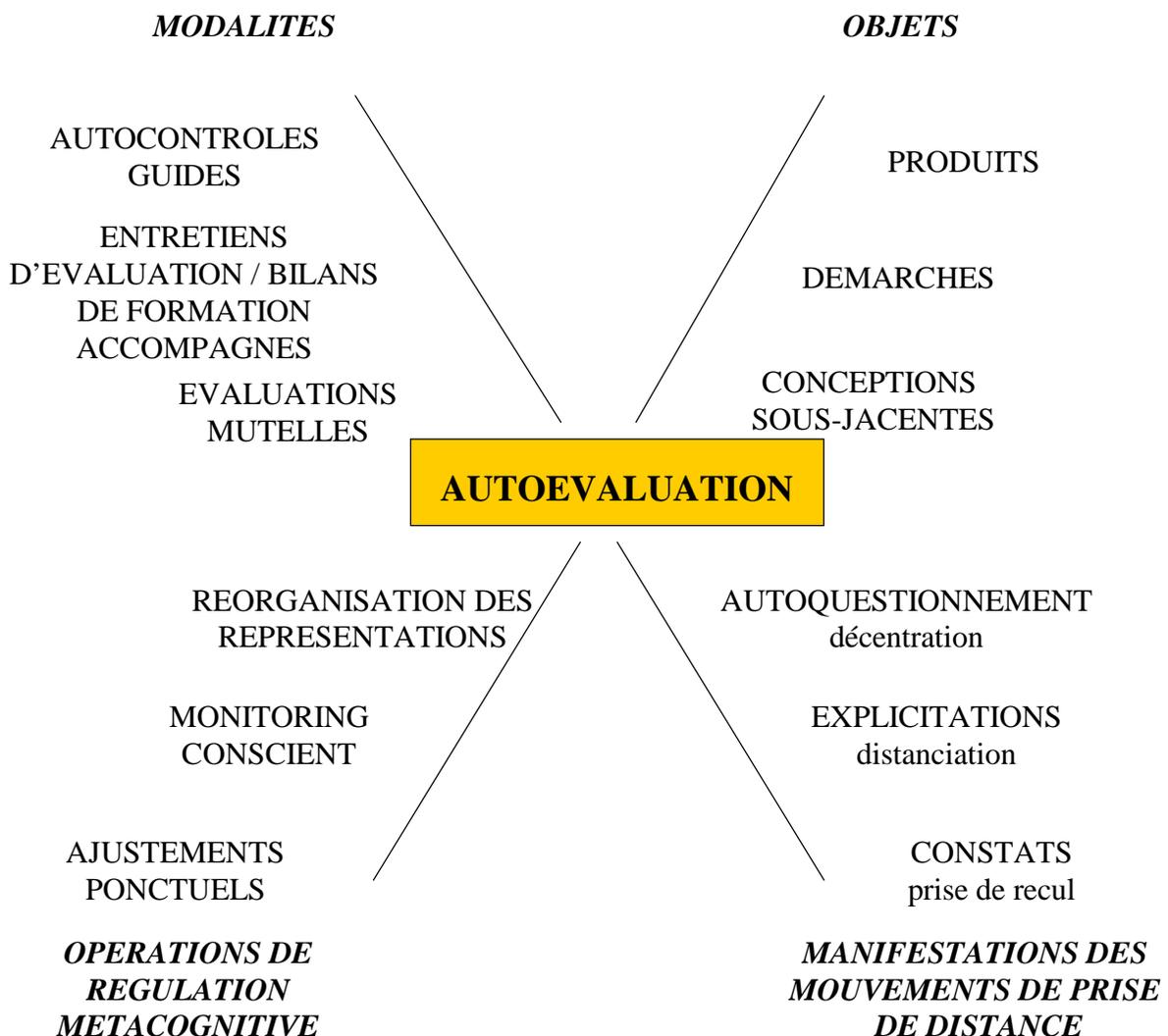


Figure 2 - Grille de lecture de l'autoévaluation

### 3.3 La régulation du système enseignement-apprentissage

L'autoévaluation et la régulation du système sont facilitées si tous les acteurs s'impliquent dans la même dynamique et se donnent les moyens d'évaluer leur action. Pour cela, il est important que chacun sache ce qu'il a à évaluer.

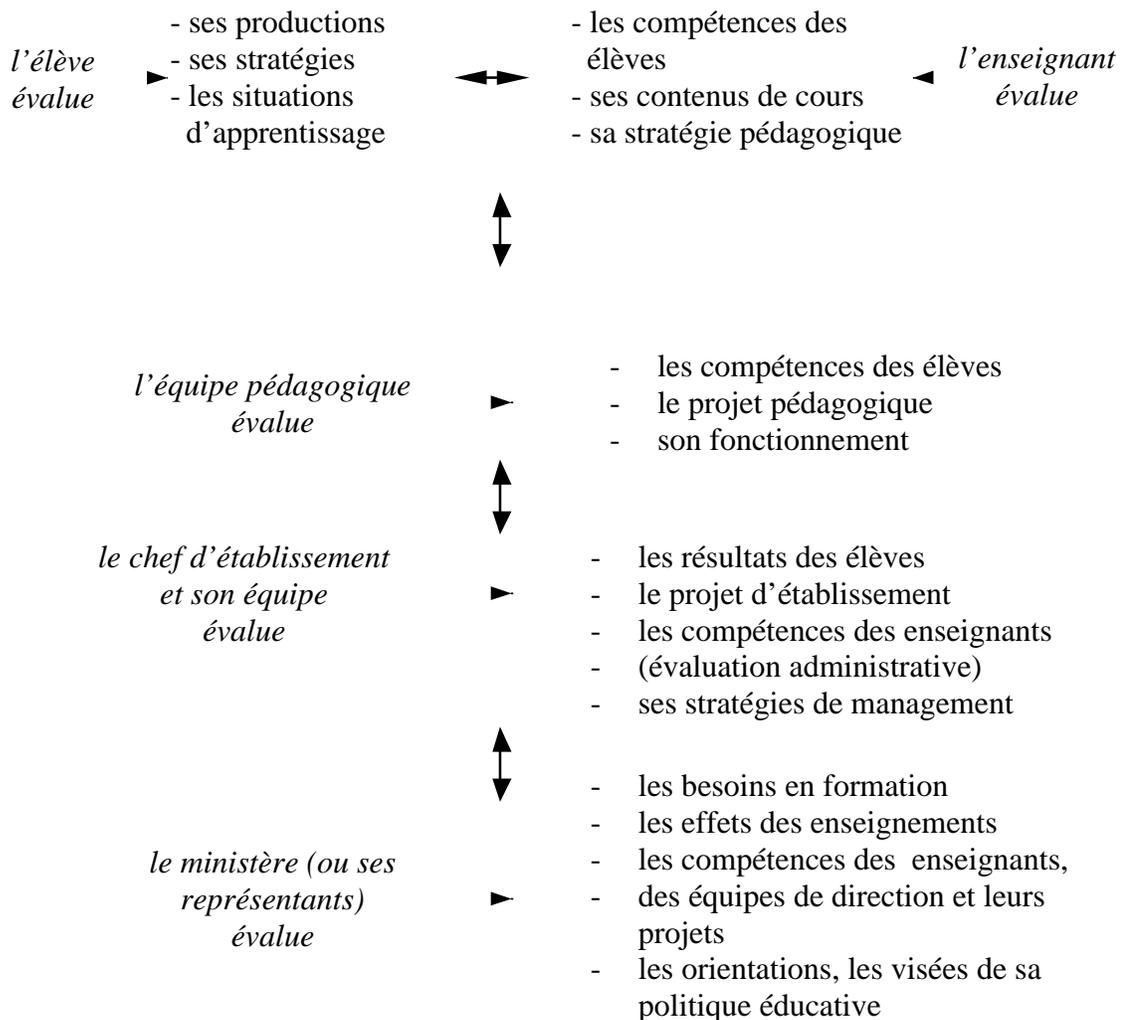


Figure 3 - Acteurs et objets de l'évaluation de l'enseignement

Un enseignant peut-il autoévaluer son enseignement autrement qu'à travers les résultats des élèves ? Oui, par exemple en conduisant une séquence avec un autre enseignant, éventuellement en questionnant les élèves en fonction de critères qui lui paraissent devoir caractériser un enseignement efficace. Autoévaluer est une démarche de lucidité, témoignant d'une responsabilisation, mais l'autoévaluation ne relève pas que du cognitif : **la personne avec ses affects est en jeu dans l'autoévaluation**<sup>9</sup>. Les jugements de valeur sur les personnes doivent en être exclus.

## **CONCLUSION : EVALUER, AUTOEVALUER, COURAGE ET MOTIVATION**

L'évaluation, ce n'est pas simple. C'est un exercice où l'on prend des risques : **on ne peut jamais être sûr d'avoir « bien » évalué**, toute évaluation renvoie des interrogations à l'évaluateur. Mais elle fait partie de la mission de tout enseignant et lui permet d'assurer ses choix. L'évaluation critériée, qui permet de sortir de l'arbitraire, est un exercice de **rigueur** et de **lucidité**. **La démarche d'évaluation formatrice, qui implique l'élève et s'intègre à son apprentissage**, est non seulement **facteur de progression mais aussi de motivation** pour l'apprenant.

---

<sup>9</sup> Cf. les pièges de l'autoévaluation recensés par Allal (1999).

## Références bibliographiques

- ALLAL (L), CARDINET (J), PERRENOUD (P) (éd.), 1979, 1981, 1983, 1985, *L'évaluation formative dans un enseignement différencié*, Berne, Peter Lang.
- ALLAL (L), 1979, Stratégies d'évaluation formative : conceptions psychopédagogiques et modalités d'application. In ALLAL & CARDINET & PERRENOUD (éd.), 1979, 1981, 1983, 1985, *L'évaluation formative dans un enseignement différencié*, Berne, Peter Lang.
- ALLAL (L), 1988, Vers un élargissement de la pédagogie de la maîtrise : processus de régulation interactive, rétroactive, proactive. In M. Huberman (sous la dir.), *Assurer la réussite des apprentissages scolaires ?* Delachaux & Niestlé, pp. 86-126.
- ALLAL (L), 1991, *Vers une pratique de l'évaluation formative, matériel de formation continue des enseignants*, Bruxelles, De Boeck.
- ALLAL (L), 1993, Régulations métacognitives : quelle place pour l'élève dans l'évaluation formative ? In L. ALLAL, D. BAIN, P. PERRENOUD, *Evaluation formative et didactique du français*, Delachaux et Niestlé, pp. 81-98.
- ALLAL (L) et Michel (Y), 1993, Autoévaluation et évaluation mutuelle en situation de production écrite. In L. ALLAL, D. BAIN, P. PERRENOUD, *Evaluation formative et didactique du français*, Delachaux et Niestlé, pp. 239-264.
- ALLAL (L), BAIN (D), PERRENOUD (P), 1993, *Evaluation formative et didactique du français*, Delachaux et Niestlé
- ALLAL (L), 1999, Impliquer l'apprenant dans le processus d'évaluation : promesses et pièges de l'autoévaluation. In C. Depover & B. Noël, *L'évaluation des compétences et des processus cognitifs*, De Boeck, pp. 35-56.
- ARDOINO (J) & BERGER (G), 1986, « L'évaluation comme interprétation », *Journal de la Formation Continue & de l'EAO*, n° 208, nov. 1986, Paris, p. 7-10 (reproduction de l'article paru dans la revue POUR, n°107, juin-juillet-août 1986, Paris, p 120-127).
- ARDOINO (J) & BERGER (G), 1989, *D'une Evaluation en miettes à UNE EVALUATION EN ACTES, le cas des universités*, Matrice, Andsha.
- ARDOINO (J), avril 1993, l'approche multiréférentielle (plurielle) des situations éducatives et formatives, *Pratiques de formation*, n°25-26, p. 15-34.
- BARBIER (JM), 1985, *L'évaluation en formation*, Paris, PUF.

- BERBAUM (J), 1991, *Développer la capacité d'apprendre*, ESF.
- BERBAUM (J), 1993, « Le développement de la capacité d'apprentissage », in J. Houssaye (sous la dir.), *La Pédagogie : une encyclopédie pour aujourd'hui*, ESF.
- BLOOM (BS), 1971, *Handbook on formative and summative evaluation of student learning*, New York, Mc Graw-Hill Book Co.
- BLOOM (BS), 1988, Le problème des deux sigmas : la recherche de méthodes d'enseignement en groupe aussi efficaces que le préceptorat. In M. Huberman (sous la dir.), *Assurer la réussite des apprentissages scolaires ?* Delachaux & Niestlé, pp. 45-83
- BONNIOL (JJ), AMIGUES (R), GUIGNARD (C), GENTHON (M), 1978, *Dispositifs d'évaluation formative, d'auto-évaluation et d'auto-correction: impact de ces dispositifs*. Rapport de recherche, MEN.
- BONNIOL (JJ), 1986, « Recherches et Formations: Pour une problématique de l'évaluation formative », in J.M. De Ketele (éd.), *L'évaluation: approche descriptive ou prescriptive ?* De Boeck, p. 119-133.
- BONNIOL (JJ), 1988, « Entre les deux logiques de l'évaluation, rupture ou continuité ? », *Bulletin de l'ADMEE*, n° 88-3, pp. 1-5.
- BONNIOL (JJ), GENTHON (M), 1989, « L'évaluation et ses critères : les critères de réalisation », *Repères*, n° 79, p. 107-115.
- CAMPANALE (F), 1995, *L'autoévaluation, facteur de transformation des conceptions et des pratiques : études d'actions de formation continue des enseignants dans le domaine de l'évaluation*, thèse pour le doctorat en Sciences de l'Education, Université Pierre-Mendès-France, Grenoble II.
- CAMPANALE (F), 1997. « Autoévaluation et transformations de pratiques pédagogiques », *Mesure et Evaluation en Education*, Vol. 20, n°1, p. 1-24.
- CAMPANALE (F), 1998, « Modélisation des processus d'autoévaluation en formation d'enseignants », *Pédagogies*, n°12, 1998.
- CAMPANALE (F), 1999, « D'un modèle de l'autoévaluation aux situations d'évaluation partagée dans la formation initiale des enseignants ». In Fabre-Cols C. Triquet E. (coord.), *Recherche(s) et Formation des Enseignants, Actes du 2ème Colloque International*, organisé par l'IUFM de Grenoble les 5, 6, 7 février 1998, IUFM académie de Grenoble, Grain d'Aile.
- CARDINET (J), 1989, « Evaluer sans juger », *Revue Française de Pédagogie*, n°88, juil., août, sept, p. 41-52.

CARDINET (J), 1992, « L'objectivité de l'évaluation », in *Formation et technologies*, n°0, commission communautaire française, 1er trimestre, p. 17-26

CHEVALLARD (Y), 1986, « Vers une analyse didactique des faits d'évaluation », in J.M. De Ketele (éd.), *L'Evaluation : approche descriptive ou prescriptive ?*, De Boeck, p. 31-59.

DE KETELE (J M), 1986, « L'évaluation du savoir-être », in J. M. De Ketele (éd.), *L'Evaluation : approche descriptive ou prescriptive ?* Bruxelles, De Boeck, p. 179-208.

GALPERINE (PI), 1980, « Formation par étapes des actions et des concepts », in N. F. Talyzina (éd.), *De l'enseignement programmé à la programmation de la connaissance*, Lille : Presses Universitaires.

GRANGEAT (M) (coor.) & MEIRIEU (P) (sous la dir.), 1997, *La métacognition, une aide au travail des élèves*, ESF.

GRANDGUILLOT (MC), 1993, *Enseigner en classe hétérogène*, Hachette-éducation

HADJI (Ch), 1989, 1990, *Evaluation, les règles du jeu*, ESF.

HADJI (Ch), 1992, *L'évaluation des actions éducatives*, PUF.

HADJI (C), 1997. *L'évaluation démythifiée*, ESF, coll. Pratiques & enjeux pédagogiques.

LE BOTERF (G), 1998, *L'ingénierie de compétences*, Editions d'Organisation.

MERLE (P), 1998, *Sociologie de l'évaluation scolaire*, PUF, Que sais-je ? n°3278.

MALGLAIVE (G), 1990, *Enseigner à des adultes*, PUF.

NOEL (B) & ROMAINVILLE (M) & WOLFS (JL), 1995, La métacognition : facettes et pertinence du concept en éducation, *Revue Française de Pédagogie*, 112 (3), 47-56.

NOIZET (G) & CAVERNI (JP), 1978, *Psychologie de l'évaluation scolaire*, PUF.

NUNZIATI (G), 1987, « Pour évaluer une formation », *Cahier de l'évaluation*, n°2, Sèvres, CIEP, p. 47-57.

NUNZIATI (G), 1990, « Pour construire un dispositif d'évaluation formatrice », *Cahiers Pédagogiques : Apprendre*, n°280 janv. 1990, pp. 48-64.

PAQUAY (L) & ALLAL (L) & LAVEAULT (D), 1990, « L'auto-évaluation en question(s) - Propos pour un débat », *Mesure et Evaluation en Education*, vol. 13, n°3, p. 5-26.

PAQUAY (L), 1994, « Prescrire l'auto-évaluation? Oui, mais... ! », *Bulletin de l'ADMEE*, n°94/2 - octobre

PIAGET (J), 1974a, *La prise de conscience*, Paris, PUF.

PIAGET (J), 1974b, *Réussir et comprendre*, Paris, PUF.

SCRIVEN (M), 1967. « The Methodology of evaluation », in Tyler (R), Gagné (RM) et Scriven (M), *Perspectives of curriculum evaluation*, Chicago, Rand Mac Nally, p. 39-83.

VERGNAUD (G), 1995, « Introduction », *Dossier : Compétences, Performances Humaines & Techniques*, n°75-76 mars-avril/mai-juin 1995, pp. 7-12.

VESLIN (O. et J.), 1992, *Corriger des copies*, Hachette-éducation

VIAL (M), 1987, *A propos du narratif... un dispositif d'évaluation formatrice en expression écrite*, Marseille : CNDP, CRDP.

VIAL (M), 1995, « Nature et fonction de l'auto-évaluation dans le dispositif de formation ». *Revue Française de Pédagogie*, n° 112, 69-76.

VIAL (M), 1997, « L'auto-évaluation comme auto-questionnement », *Les Cahiers d'Aix*, n°12, Université de Provence, Département des Sciences de l'Éducation.

## EVALUATION - QUELQUES DEFINITIONS DE CONCEPTS-CLES

**EVALUER**, c'est **mettre en relation** des éléments issus d'un observable ou **référé** et un **réfèrent** pour **produire de l'information éclairante** sur l'observable, afin de prendre des décisions (HADJI Ch., 1990, *Evaluation, les règles du jeu*, ESF).

**le réfèrent** : « Ce à quoi se rapporte pour devenir plus intelligible un matériel donné » (ARDOINO J. & BERGER G., 1989, D'une Evaluation en miettes à une évaluation en actes, Matrice-ANDSHA). Il comporte l'objectif de l'action (les compétences visées) et les critères sur lesquels on s'appuiera pour évaluer l'observable, c'est à dire la production de l'apprenant en fonction de la tâche prescrite. Si on évalue un produit fini, les critères seront des qualités attendues de ce produit ; s'il on évalue un processus, les critères seront des opérations invariantes à réaliser pour fabriquer ce produit et qui mettent en jeu les compétences visées.

**le référé**, c'est l'observable (la production / la démarche) et les indicateurs qui permettront de se prononcer en fonction des critères.

Cf. Compétence, Tâche, Critères d'évaluation, Indicateurs

A la suite des travaux d'Ardoino et Berger, on distingue contrôle et évaluation.

**LE CONTROLE** mesure des écarts *entre* des produits, des démarches et une *norme extérieure, préétablie*. *Le contrôle est monoréférentiel*. Il suppose un avant, un après. Il est régi par des critères de conformité, de logique, de cohérence. Il vérifie pour valider ou rejeter, corriger ou sanctionner. Il vise à normaliser.

**L'EVALUATION**, pluri-référentielle, englobe et dépasse le contrôle. Elle privilégie le qualitatif sur la quantitatif. Au delà du contrôle analytique, elle est conçue comme un processus intervenant dans un système ouvert, en évolution, en vue d'en élucider le fonctionnement et l'évolution. Les référentiels s'élaborent dans le processus même de l'évaluation. En ce sens, elle ne peut être que partagée par les acteurs.

Saisir la réalité à travers des procédures de contrôle ou la saisir plus largement à travers une évaluation qui cherche à comprendre sont des attitudes révélatrices d'un système explicatif du Monde qui met en jeu les valeurs profondes du sujet

**EVALUATION SOMMATIVE** : « évaluation par laquelle on fait un inventaire des compétences acquises, ou un bilan, après une séquence de formation d'une durée plus ou moins longue » (HADJI Ch., 1990, *Evaluation, les règles du jeu*, ESF). L'évaluation sommative met l'accent sur les performances (les productions réalisées) évaluées en fonction des critères de réussite. Elle relève davantage du contrôle que de la régulation, sans l'exclure pour autant.

**EVALUATION FORMATIVE** : « évaluation dont l'ambition est de contribuer à la formation (HADJI Ch., 1990, *Evaluation, les règles du jeu*, ESF). Elle privilégie la régulation en cours de formation. Elle tente de fournir à l'apprenant des informations pertinentes pour qu'il régule ses apprentissages et elle renvoie à l'enseignant un feedback sur son action qui lui permet d'adapter son dispositif d'enseignement. L'évaluation formative met l'accent davantage sur les processus évalués à travers les critères de réalisation.

### **COMPETENCE**

On peut la définir comme un ensemble de savoir-faire conceptualisés dont la maîtrise implique la mise en œuvre combinée de savoirs formalisés (connaissances scientifiques et techniques), de savoirs pratiques et comportementaux, d'opérations mentales (MALGLAIVE G., 1990, *Enseigner à des adultes*, PUF ; VERGNAUD G., 1995, " Introduction ", Dossier : Compétences, *Performances Humaines & Techniques*, n°75-76 mars-avril/mai-juin 1995, pp. 7-12). La compétence, au delà de cet ensemble de savoirs et de savoir-faire de différents types à mobiliser, combiner, transposer, suppose un « savoir agir dans une situation professionnelle [ou scolaire] complexe en vue d'une finalité » (LE BOTERF G., 1998, *l'ingénierie de compétences*, Editions d'Organisation). La compétence se développe et s'évalue à travers la réalisation de tâches et des effets qui en résultent (= la performance). On peut distinguer compétences générales ou transversales (dégager l'essentiel, problématiser, travailler en équipe...) et des compétences spécifiques à une discipline (rendre compte d'une expérience en SVT).

**TACHE** (= but, environnement, opérations)

C'est l'activité prescrite à un sujet en vue d'atteindre un but (produire tel objet observable), dans un environnement donné au moyen d'actions ou d'opérations. La tâche gagne à être définie par des critères qui définissent l'attendu par rapport à un ou plusieurs objectifs. Elle suppose des consignes qui finalisent l'activité (le but / la production demandée), indiquent éventuellement des éléments de méthode, et les conditions de réalisation. La réalisation de la tâche nécessite la mise en œuvre d'opérations invariantes qui lui sont spécifiques et que le sujet effectue en fonction de démarches qui lui sont propres.

**EVALUATION NORMATIVE** : une évaluation qui situe les individus les uns par rapport aux autres, en fonction des scores obtenus par les membres d'un groupe de référence (HADJI Ch., 1990, *Evaluation, les règles du jeu*, ESF). Dans l'évaluation scolaire courante, la norme de référence est bien souvent constituée par les performances moyennes du groupe classe.

**EVALUATION CRITERIEE** : « évaluation qui apprécie un comportement en le situant par rapport à une cible (le critère qui correspond à l'objectif à atteindre) » (HADJI Ch., 1997. *L'Evaluation démystifiée*, ESF).

## CRITERES D'EVALUATION

Dimensions abstraites, qualitatives d'un objet sur lesquelles on s'appuie pour se prononcer sur l'objet. On distingue critères de réussite et critères de réalisation (BONNIOL JJ & GENTHON M., 1989, « L'évaluation et ses critères: les critères de réalisation », *Repères*, n° 79, p. 107-115) :

- **les critères de réussite** sont des caractéristiques du produit attendu retenues pour l'estimer et l'apprécier en fonction d'un référent pré-établi (conformité à la consigne, exactitude / complétude, cohérence, pertinence, rigueur, communicabilité, originalité, efficacité...)
- **les critères de réalisation** ou opérations invariantes impliquées par la réalisation de la tâche découlent d'une analyse de la tâche (ex. pour le résumé : identifier le type de texte à résumer, la fonction du résumé et le destinataire, distinguer idées et illustrations de ces idées, dégager la structure du texte, sélectionner les informations / idées événements essentiels, les reformuler, organiser le résumé en reprenant la structure et les marques d'énonciation (suivant la fonction du résumé), rédiger l'ensemble de façon personnelle et concise. Ils servent aux apprenants pour se guider dans la réalisation et améliorer leur action, autrement dit pour la réguler et aux enseignants pour repérer où se situent les erreurs des élèves et produire un feedback en cours d'apprentissage.

## INDICATEURS

Un indicateur est la version ou conversion concrète d'un critère, un indice pris dans l'observable qui permet de dire si l'objet répond au critère (ex. les idées présentes dans le résumé, l'enchaînement de ces idées ; le pourcentage de réussite des étudiants aux tests de contrôle ...). Plusieurs indicateurs, se combinant, peuvent correspondre à un même critère.

## L'AUTOEVALUATION

Il s'agit d'une évaluation interne conduite par le sujet de sa propre action et de ce qu'elle produit. C'est un processus d'altération de son référentiel d'action au cours de confrontations entre son propre référentiel et celui ou ceux d'autrui (CAMPANALE F., 1997. « Autoévaluation et transformations de pratiques pédagogiques ». *Mesure et Evaluation en Education*, Vol. 20, n°1, p. 1-24).

Elle se traduit par un dialogue de soi à soi, une réflexion métacognitive (ALLAL L. & BAIN D., PERRENOUD P., 1993, *Evaluation formative et didactique du français*, Delachaux et Niestlé). L'autoévaluation délibérée, qui amène le sujet à interroger, réguler, transformer son action (produit, démarche, conceptions implicites sous-jacentes) ne peut être imposée (PAQUAY L., 1994, « Prescrire l'auto-évaluation? Oui, mais... ! », *Bulletin de l'Admee*, n°94/2 - octobre).