

PERFORMA

L'évaluation formative

planification et instrumentation

ROBERT GAUTHIER
MICHEL SAINT-ONGE

MODULE 1

DEFINITIONS

dossier de l'animateur



DE SOREL-TRACY
DE LA MAIRIE, TRACY, QUÉ J3R 5B9

785616
v.1

Faculté
d'éducation

UNIVERSITÉ
DE SHERBROOKE



Centre de documentation collégiale
1111, rue Lapierre
Lasalle (Québec)
H8N 2J4

PERFORMA

CENTRE DE
DOCUMENTATION COLLÉGIALE

15 AOUT 2005

L'évaluation formative

planification et instrumentation

ROBERT GAUTHIER
MICHEL SAINT-ONGE

DOSSIER DE L'ANIMATEUR

Centre de documentation collégiale
1111, rue Labrière
Laval (Québec)
H3N 2Y4

Centre de
documentation collégiale
1111, rue Labrière
Laval (Québec)
H3N 2Y4

Centre de documentation collégiale
1111, rue Labrière
Laval (Québec)
H3N 2Y4

23824
785616
V-1

Centre de documentation collégiale

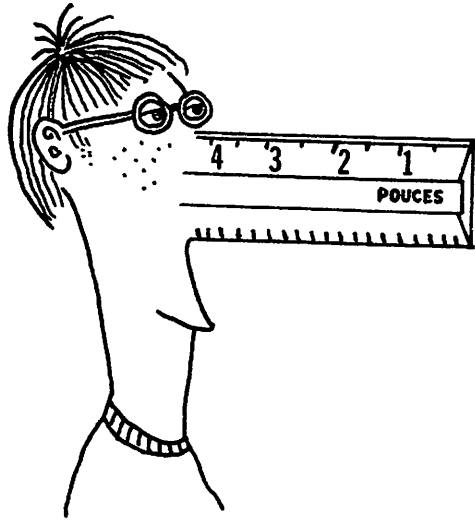
LES OBJECTIFS D'APPRENTISSAGE DE CE MODULE:

1. Définir mesure, évaluations formative et sommative, normative et critériée, évaluation diagnostique.
2. Une situation étant décrite, identifier le type d'évaluation pertinent.

DEROULEMENT:

1. Exposé (à l'aide d'aéétates)
 2. Activité d'évaluation formative
 3. Retour sur les définition à partir des résultats de l'évaluation formative.
-

M E S U R E



Toute activité qui vise à recueillir des résultats ou autres indices relatifs à diverses performances ou opérations.

La mesure exige l'usage d'un instrument (règle, thermomètre, test...) pour identifier la quantité et/ou la qualité d'une propriété ou d'un comportement.



E V A L U A T I O N



Toute activité qui vise à analyser et à INTERPRETER des résultats ou des indices provenant de la mesure afin de prendre de meilleures décisions.

L'évaluation est la prise de décision sur l'adéquation, la bonté, la satisfaction de la propriété ou du comportement mesuré. Elle devrait par conséquent s'appuyer sur les résultats d'une mesure précise et consciencieuse.

QU'EST-CE QU'ÉVALUER ?

Prendre une décision, par rapport à un critère, sur l'adéquation de l'atteinte d'objectifs d'apprentissage de capacités déterminées.

Cette décision résulte de l'interprétation de résultats de la mesure obtenus grâce à des tests objectifs et congruents, en fonction du contenu des épreuves et en tenant compte des attitudes, des intérêts et des valeurs de l'apprenant.

EVALUATION

NORMATIVE

Lorsque l'évaluation vise à situer la position d'un objet dans sa catégorie, en ce qui concerne la caractéristique envisagée.

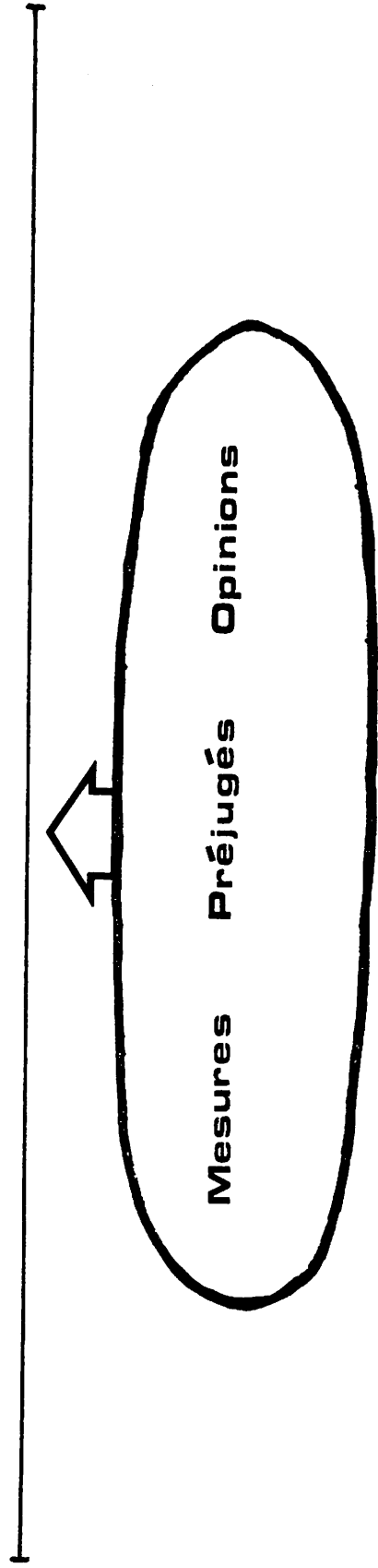
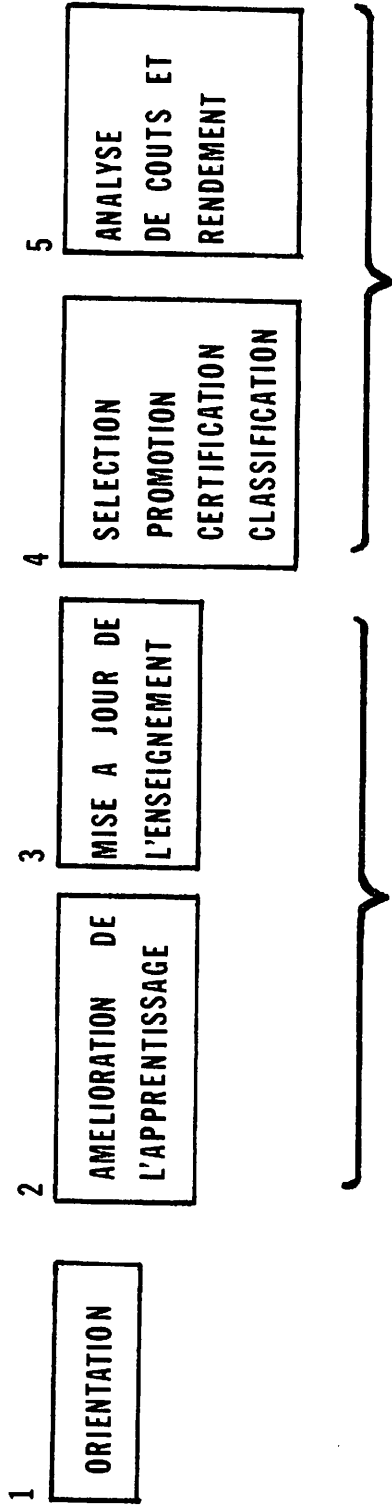
L'objet est comparé aux autres objets de sa classe.

CRITERIEE

Lorsque l'évaluation vise à déterminer si l'objet est satisfaisant par rapport à un critère, indépendamment de sa position dans sa classe.

L'objet n'est pas comparé aux autres objets de sa classe mais à un critère.

RÔLES DE L'ÉVALUATION
(cinq types principaux de décisions)



EVALUATION

formative



Un feedback diagnostique pour l'élève et pour le maître portant sur une partie d'une activité d'apprentissage et permettant une action corrective si cela s'impose.

BUT: Améliorer l'apprentissage.

Reconnaître où et en quoi un élève éprouve une difficulté et l'en informer.

FORME: Cette évaluation ne se traduit pas en un verdict de réussite ou d'échec.

sommative



Une évaluation portant sur l'ensemble ou une partie importante d'un cours.

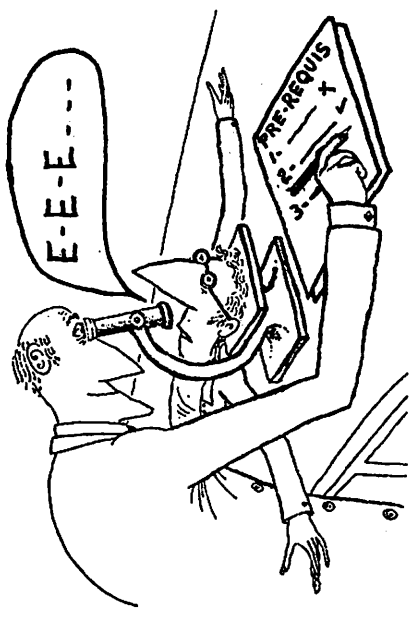
BUT: Certifier l'apprentissage.

Reconnaître si un élève a atteint ou non les objectifs d'un cours.

FORME: Cette évaluation se traduit par un verdict de réussite ou d'échec.

EVALUATION DIAGNOSTIQUE :

CET ELEVE EST-IL
VRAIMENT PRET A ENTREPRENDRE
UNE NOUVELLE EXPERIENCE
D'APPRENTISSAGE ???



INTERVENTION D' EVALUATION

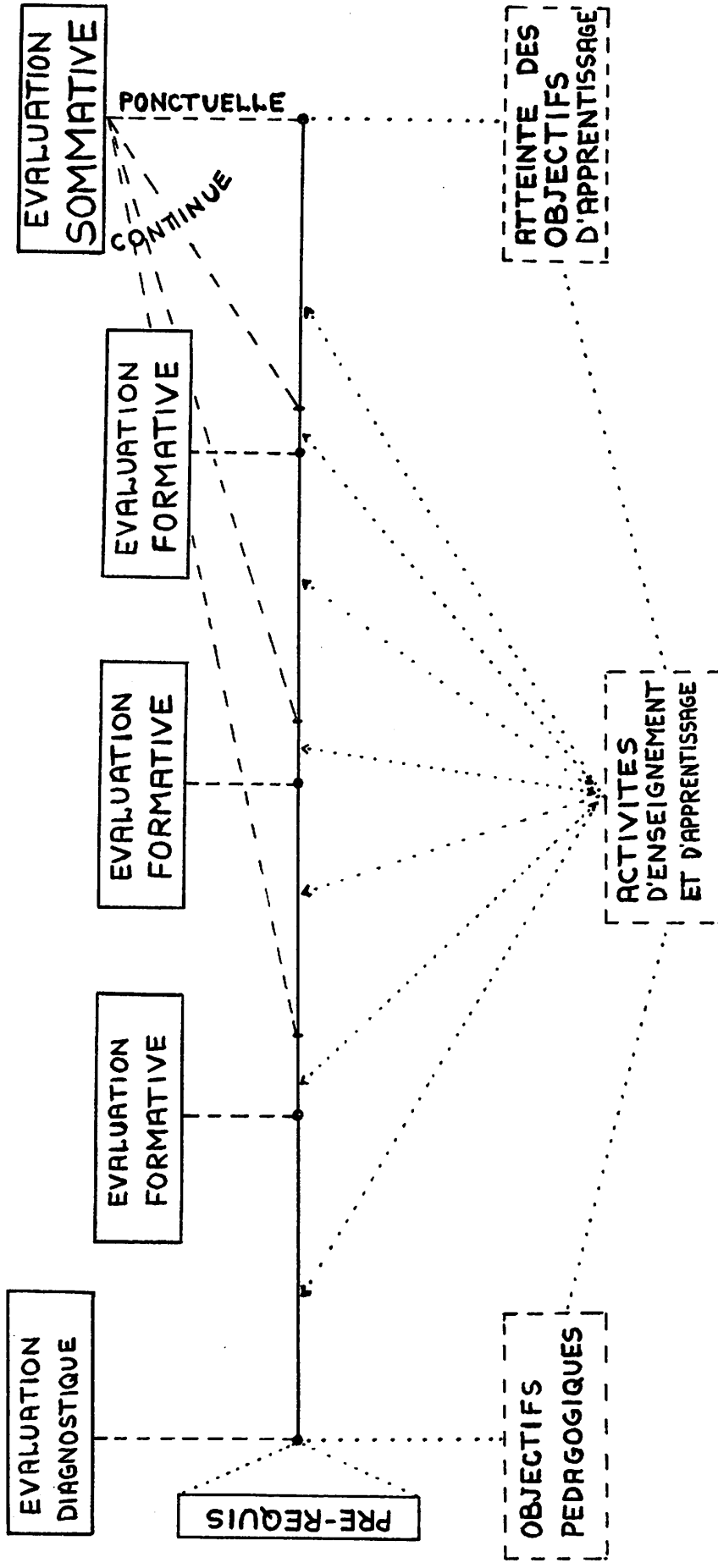
DESTINEE A INFORMER L'ELEVE

ET L'ENSEIGNANT SUR LA

MAITRISE D'HABILETES PRE-REQUISES

POUR L'ATTEINTE DES OBJECTIFS D'UN COURS DONNE.

FORMES D'EVALUATION



L'évaluation dans un système d'enseignement

<i>But du système (idéologie partagée)</i>	<i>Forme d'enseignement</i>	<i>Fonction de la régulation</i>	<i>Formes d'évaluation</i>	<i>Moment</i>
SELECTIONNER les élèves qui maîtrisent le plus grand nombre d'objectifs pédagogiques	Enseignement .collectif .indifférencié	Assurer que les caractéristiques des élèves répondent aux exigences du système.	<u>Sommative</u> (information en vue de décider de la certification)	A la fin d'une période de formation
Amener TOUS les élèves à la maîtrise de certains objectifs pédagogiques	Enseignement .individualisé .différencié	Assurer que les moyens de formation correspondent aux caractéristiques des élèves.	<u>Pronostique</u> (information en vue de décider de l'orientation)	Au début d'un cycle de formation
			<u>Formative</u> (information en vue de guider l'apprentissage)	Pendant la période de formation

Documents:

- SCALLON, Gérard, L'évaluation des étudiants et les principales conceptions de la mesure et de l'évaluation dans PROSPECTIVES, vol. 10, n^o 4, octobre 1974, Montréal, pp. 263-269
- FONTAINE, France, Les modèles d'évaluation de l'apprentissage

l'évaluation des étudiants et les principales conceptions de la mesure et de l'évaluation

par Gérard SCALLON *

L'évaluation des étudiants n'est pas une préoccupation nouvelle au Québec. Depuis quelques années cependant, elle est devenue un point névralgique dans de nombreuses discussions ou remises en question. Alors qu'elles étaient dans le passé sous la responsabilité presque exclusive des administrateurs, les activités d'évaluation exercées par les enseignants et les étudiants prennent de plus en plus d'importance. Mais ce n'est qu'une tendance parmi plusieurs, car certaines conceptions de l'enseignement remettent parfois en question le bien-fondé même de l'évaluation des étudiants. Dans ce débat, le profane, l'homme de la rue et même plusieurs professionnels de l'éducation éprouvent une certaine confusion.

Dans le domaine de l'évaluation des étudiants, la mise en marché de concepts, comme la mesure normative, la mesure critériée, l'évaluation « formative », l'évaluation « sommative » ou l'évaluation continue, ne se réalise pas sans problème. Ces concepts ne sont pas d'hier mais leur utilisation est nouvelle, et leur injection dans notre système d'éducation risque d'augmenter la confusion si l'on ne se préoccupe pas de bien les clarifier au point de départ.

Nous nous proposons dans cet article d'établir quelques distinctions qui apparaissent à la base de toute discussion au sujet du bien-fondé de

l'évaluation des étudiants et qui semblent nécessaires à la compréhension de plusieurs termes utilisés. Nous n'avons pas la prétention de présenter ici un texte qui soit savamment documenté et qui traite adéquatement des problèmes que soulève l'évaluation des étudiants en milieu scolaire. Notre préoccupation sera plutôt de définir le plus simplement possible un certain nombre de termes qui sont de plus en plus utilisés dans les écrits qui traitent de l'évaluation des étudiants.

Mesurer et évaluer

Une distinction qui nous apparaît fondamentale est celle que nous devons établir entre *mesurer* et *évaluer*. Nombre de théoriciens de l'évaluation s'entendent pour affirmer que mesurer est une opération qui consiste essentiellement à associer des symboles à des objets, à des événements ou à des personnes selon des règles précises. Les symboles qui ont été utilisés au cours de cette opération s'appellent, une fois l'opération terminée, les résultats de la mesure. Dans la plupart des cas, ces symboles sont des nombres qui pour

* L'auteur est professeur en mesure et évaluation de la faculté des Sciences de l'éducation de l'Université Laval.

l'instant sont purement descriptifs. Ils doivent leur signification à l'unité de mesure qui les accompagne. Ainsi, on peut associer le nombre 7 à une table parce que cette table a une longueur de 7 pieds. Dans ce cas, le pied, qui est l'unité de mesure utilisée, a été appliqué un certain nombre de fois sur la longueur de la table. Le nombre 7 prend donc sa signification en vertu de sa relation à une longueur standard qui est le pied. La plupart des résultats de mesure obtenus en éducation n'ont pas cette propriété, à cause de la difficulté de trouver une unité de mesure pouvant servir de standard. Par exemple, un résultat comme « 40 bonnes réponses » à une épreuve de *rendement* est difficile à interpréter, parce que les 40 questions correctement résolues pour obtenir ce résultat ne constituent pas nécessairement des unités égales.

Évaluer, d'autre part, est une opération qui consiste à porter des jugements de valeur ou à accorder des valeurs à des objets, à des événements ou à des personnes en comparant ces objets, événements ou personnes avec un critère quelconque. Cette opération peut aussi s'appliquer aux résultats d'une mesure. Une fois l'opération terminée, on donne le nom d'évaluation aux jugements de valeur qui ont été portés. Des expressions comme « excellent », « médiocre », « supérieur à la moyenne », « succès » ou « échec » etc. sont des exemples d'évaluations rencontrées en éducation. Il faut comprendre que l'évaluation d'un résultat de mesure est d'autant plus importante que ce résultat est difficile à interpréter, à cause de l'absence d'une unité de mesure pouvant lui donner une signification.

Quelques exemples nous feront mieux saisir la distinction que nous venons d'établir entre mesurer et évaluer. Quelqu'un nous informe qu'un de ses amis est âgé de 90 ans. L'expression *90 ans* (ou le nombre 90) est le résultat d'une mesure. Pour associer le nombre 90 à l'individu concerné dans notre exemple, on a sans doute appliqué une règle précise qui consiste à déterminer combien de fois une certaine unité de mesure, l'année d'existence, est contenue dans l'écart qui sépare deux moments précis: la date de naissance de l'individu et la date où nous nous trouvons présentement. Si notre interlocuteur nous avait dit qu'un de ses amis est très âgé, il nous aurait alors communiqué une évaluation. Voici un autre exemple. Une personne nous apprend qu'elle a rencontré quelqu'un dont la hauteur physique est de 8 pieds. La règle qui

a été utilisée ici pour associer le nombre 8 à un individu nous est connue. Si cette même personne nous avait dit qu'elle a rencontré un géant, elle nous aurait communiqué alors son propre jugement de valeur, c'est-à-dire son évaluation de la grandeur physique de l'individu qu'elle a rencontré. On peut voir aisément, avec les deux exemples qui précèdent, que le résultat de la mesure est purement descriptif et ce sans qu'aucun jugement de valeur ne soit porté intentionnellement. Il s'agit là de l'une des principales caractéristiques de la mesure qui distingue celle-ci de l'évaluation. On objectera que la distinction n'est pas très claire dans les deux exemples, car tout le monde sait qu'un homme âgé de 90 ans est très âgé et qu'un individu mesurant 8 pieds de hauteur est un géant. C'est que nous faisons face ici à une situation très fréquente dans la vie courante, et dans laquelle le résultat d'une mesure s'inscrit facilement dans un cadre de référence qui correspond à notre expérience vécue. Nous connaissons tous l'unité de mesure qui a été utilisée dans chaque exemple: l'année d'existence dans le cas de l'âge et le pied dans le cas de la grandeur physique. De plus, nous connaissons les possibilités actuelles de la longévité humaine de même que la répartition de la grandeur physique chez un très grand nombre d'individus. Ce sont ces deux informations, l'une concernant l'unité de mesure utilisée et l'autre concernant les possibilités ou les limites du phénomène étudié, qui nous permettent très souvent de porter notre propre jugement de valeur à partir uniquement d'un résultat de mesure. Notons que dans les deux exemples qui précèdent, les jugements de valeur « très âgé » et « géant » sont des évaluations relatives à un ensemble d'individus dont on connaît la longévité et la grandeur physique. Le critère qui a servi de base à de telles évaluations est la caractéristique d'un ensemble d'individus. Il suffit que cette caractéristique soit modifiée pour que l'évaluation qui lui est relative soit elle aussi modifiée. Par exemple, si l'on parvenait un jour à prolonger la vie humaine jusqu'à 200 ans, alors le jugement de valeur « très âgé » ne conviendrait plus pour une personne de 90 ans.

Dans certains contextes, l'unité de mesure et les possibilités du phénomène étudié ne font pas partie de notre expérience vécue. Il devient alors difficile, sinon impossible, d'évaluer nous-mêmes à la suite du résultat d'une mesure. Prenons une situation fictive comme exemple. Un astronaute revient d'un voyage interplanétaire et nous

apprend que le premier habitant qu'il a rencontré sur la planète BX-12 mesurait 3 CVZ de hauteur (le CVZ étant l'unité de longueur utilisée par les habitants de la planète BX-12). Nous sommes en face du résultat d'une mesure et il nous est impossible d'évaluer la grandeur de cet habitant ou de porter un jugement de valeur comme: « très petit », « de grandeur moyenne » ou « très grand ». C'est que nous ne possédons ni l'une ni l'autre des deux informations qui sont essentielles à l'évaluation. D'abord, nous ne connaissons pas le CVZ et ensuite, nous n'avons pas la moindre idée de la grandeur physique de l'ensemble des habitants de la planète BX-12. Supposons que notre astronaute nous fournit des indices quant à cette deuxième information: il nous dit que la grandeur physique des habitants de cette planète varie entre 3CVZ et 9 CVZ. Déjà ce renseignement additionnel nous permet de déduire que le premier habitant qu'il a rencontré, avec ses 3 CVZ, se situe fort probablement parmi les plus petits habitants de la planète BX-12. Il faut noter que cette évaluation est limitée au cadre, de référence défini par la grandeur physique des habitants de la planète concernée. Si nous devions évaluer la grandeur de notre habitant en fonction de nos dimensions humaines, il nous faudrait connaître l'unité de mesure, le CVZ. Supposons que le CVZ équivaut approximativement à trois pieds de longueur. Nous sommes capables maintenant de déduire que l'un des plus petits habitants de la planète BX-12 serait considéré comme un géant sur notre terre. Ces deux jugements de valeur n'ont rien de contradictoire et illustrent le caractère relatif de l'évaluation.

Le contexte que nous venons de décrire, dans lequel nous ne connaissons pas l'unité de mesure et les possibilités du phénomène étudié, est très fréquent en éducation. C'est le problème que nous devons résoudre lorsqu'il s'agit d'évaluer l'intelligence d'une personne, ses aptitudes, ses intérêts ou la qualité de son rendement scolaire à partir d'un résultat de mesure. Le plus souvent, ce résultat de mesure est un nombre qui indique une quantité de bonnes réponses à un ensemble de questions posées dans un test ou dans une épreuve. Ainsi, un résultat de 60 obtenu par un étudiant à une épreuve de rendement, ne nous indique rien de la qualité de cet étudiant ou de sa compétence. Pour porter un jugement de valeur, il nous faudrait connaître l'une ou l'autre des deux choses. Si nous étions familiers avec la nature des problèmes que l'étudiant a pu résoudre dans cette épreuve, nous pourrions

alors inférer les capacités ou les habiletés dont cet étudiant peut faire preuve. Il s'agit ici d'une évaluation qui fait abstraction des résultats obtenus par les autres étudiants. De plus, notre connaissance de la nature des problèmes résolus ou du contenu de l'épreuve est un compromis face à la difficulté d'identifier une unité de mesure ou un standard. Si, d'autre part, nous connaissions les possibilités de résultats à cette épreuve, telles qu'elles ont été révélées par les résultats obtenus par d'autres étudiants, nous pourrions alors porter un jugement de valeur comme « excellent », « médiocre », « supérieur à la moyenne » ou « normal ». Il s'agit, dans ce deuxième cas, d'une évaluation dont le critère de base est la caractéristique d'un ensemble d'étudiants.

Le critère pouvant servir de base à l'évaluation d'un étudiant est une notion fondamentale dans la distinction qui sera établie entre la mesure critériée et la mesure normative. Nous y reviendrons un peu plus loin dans cet article.

Une définition fonctionnelle de l'évaluation

Nous avons défini l'évaluation comme une opération distincte de la mesure. Cette définition est incomplète, car elle ne met pas en évidence le rôle de l'évaluation. On ne porte pas de jugements de valeur comme « excellent » ou « inférieur à la moyenne » uniquement pour le simple plaisir d'évaluer ou pour cataloguer. Malheureusement, trop de textes sur le sujet entretiennent la fausse perception que l'évaluation est une activité suffisante en elle-même et ouvrent ainsi la porte à des discussions non justifiées. Il existe plusieurs définitions de l'évaluation qui nous permettent d'échapper à cette difficulté. Toutes ces définitions ont un caractère commun. L'évaluation y est vue comme un ensemble de procédés qui consistent à identifier la nature de certaines décisions, à recueillir les informations appropriées en vue de communiquer des données utiles à ceux qui devront prendre ces décisions.

Les décisions qui sont prises en éducation sont trop nombreuses pour en dresser une liste exhaustive. Le choix d'un procédé d'enseignement, l'implantation et le maintien d'un régime pédagogique particulier à un niveau quelconque, l'admission des étudiants dans une institution

d'enseignement, l'inscription à un programme d'études et l'organisation de rythmes particuliers d'apprentissage pour des étudiants déficients sont des exemples variés de prises de décision. Les nombreuses décisions qui concernent chaque étudiant peuvent être prises par plusieurs personnes: enseignants, parents, administrateurs, conseillers d'orientation, étudiants, employeurs etc. Ces décisions, pour être prises de façon éclairée, doivent être précédées d'une évaluation appropriée. Le concept de prise de décision est donc fondamental dans le domaine de l'évaluation pour deux raisons importantes. Premièrement, il nous oblige à réaliser que l'évaluation n'est pas une fin en elle-même mais un moyen qui est au service d'une prise de décision. Deuxièmement, les divers modes d'évaluation que nous rencontrons de nos jours se définissent par le contexte qui caractérise certaines prises de décision. C'est ainsi que nous distinguerons la mesure normative de la mesure critériée, et l'évaluation sommative de l'évaluation formative.

Mesure normative & mesure critériée

Dans le domaine de l'apprentissage et du rendement scolaire, il existe deux critères de comparaison qui peuvent servir de base à l'évaluation des étudiants: (1) la caractéristique d'un groupe d'étudiants et (2), un seuil de rendement ou de performance. Voyons par un exemple comment chacun de ces deux critères de comparaison peut être appliqué à l'évaluation d'un résultat de mesure. Un étudiant a obtenu un résultat de 65 à une épreuve de rendement. Si le critère de comparaison adopté pour évaluer cet étudiant était l'ensemble des résultats obtenus à la même épreuve par ses confrères de classe, nous pourrions constater par exemple qu'un résultat de 65 se situe parmi les 10 pour cent de résultats les plus faibles qui ont été observés dans la classe. Nous pourrions alors évaluer cet étudiant comme étant nettement inférieur à la moyenne et c'est un jugement de valeur basé sur la caractéristique du groupe auquel cet étudiant appartient. Si, d'autre part, le critère de comparaison pour évaluer l'étudiant qui a obtenu un résultat de 65 était un seuil de performance correspondant à un résultat de 60 par exemple, nous pourrions alors affirmer que l'étudiant a réussi l'épreuve peu importe les résultats d'autres étudiants à cette épreuve. Dans ce deuxième cas,

la notion de réussite est une évaluation basée sur un seuil de performance (ou critère de performance).

L'évaluation en fonction d'une caractéristique de groupe est dite « normative », car le plus souvent cette caractéristique s'exprime par la moyenne des résultats obtenus par un ensemble d'étudiants à une épreuve de rendement. La moyenne constitue alors une norme avec laquelle le résultat de chaque étudiant est comparé. L'évaluation en fonction d'un seuil ou critère de performance est dite critériée. Les qualificatifs « normative » et « critériée » ne sauraient être appliqués à la mesure comme telle, celle-ci étant exempte de tout jugement de valeur. Pourtant, l'emploi d'expressions comme « mesure normative » (norm-referenced testing) et « mesure critériée » (criterion-referenced testing), s'est vite généralisé dans les textes qui ont été publiés sur le sujet. Il faut comprendre ici que l'expression « mesure normative » signifie un type d'instrument de mesure (test, examen ou épreuve) utilisé en vue d'une évaluation normative, et que l'expression « mesure critériée » signifie un type d'instrument de mesure utilisé en vue d'une évaluation critériée. Cette interprétation est justifiée par le fait qu'il est impossible de déterminer, uniquement à partir de l'apparence physique des questions qui le composent, si un test est normatif ou critérié. Pour ce faire, il faut absolument connaître le mode d'évaluation pour lequel ce test a été conçu.

La différence entre la mesure normative et la mesure critériée ne se situe pas uniquement au niveau du mode d'évaluation utilisé. Le mode de communication des résultats de mesure, la rédaction des questions et le contexte de la prise de décision sont également d'autres aspects pour lesquels la mesure normative et la mesure critériée sont nettement différentes.

Dans le cas de la mesure normative, on se préoccupe surtout de communiquer des résultats qui reflètent le mieux possible la position relative occupée par chaque étudiant dans un groupe donné, qu'il s'agisse d'une classe, de plusieurs classes ou d'un ensemble d'étudiants ayant subi le même examen. C'est pourquoi la note initiale obtenue par chaque étudiant à un examen est convertie en rang centile, en rang stanine, en rang cinquième ou en score-type. Dans plusieurs institutions d'enseignement, la note est convertie en lettres A, B, C, D ou E selon la région autour

de la moyenne à laquelle le résultat de l'étudiant appartient.

Dans le cas de la mesure critériée, l'accent est mis sur la description des habiletés ou capacités maîtrisées par chaque étudiant, en vue de faciliter une évaluation en fonction d'un seuil de performance. Avec une épreuve composée de questions mesurant la même habileté, on rapportera la note initiale de chaque étudiant tout en indiquant le seuil de performance exigé. Avec une épreuve mesurant des habiletés différentes, on tentera de décrire le mieux possible le niveau de compétence de chaque étudiant en regard de chaque habileté mesurée. La communication des résultats de la mesure critériée n'implique pas nécessairement des informations quantitatives et peut revêtir des aspects qualitatifs. On peut concevoir, par exemple, un bulletin d'information contenant en énumération les habiletés maîtrisées par un étudiant à la suite d'une séquence d'apprentissage. Malgré les diverses formes que peut revêtir la communication des résultats, il est important de souligner que dans la perspective d'une évaluation critériée, l'évaluation de chaque étudiant est totalement indépendante de l'évaluation des autres étudiants.

Du côté de la rédaction des questions et de la construction d'épreuves de rendement, la différence entre la mesure normative et la mesure critériée se manifeste par une différence dans l'importance accordée aux différences individuelles. Dans le cas de la mesure normative, les questions sont rédigées en vue d'accentuer les différences qui peuvent exister entre les étudiants. Ce but est atteint à la limite lorsque parmi les résultats obtenus, il n'existe pas deux résultats qui sont identiques. Dans le cas de la mesure critériée, la recherche de différences individuelles n'est pas une préoccupation majeure. Dans le cas de la mesure normative comme dans le cas de la mesure critériée, la rédaction de questions d'examen doit s'inspirer des objectifs d'un programme ou d'un cours. Cependant, en mesure normative, la recherche de différences individuelles amène parfois des concessions importantes au niveau de la rédaction de questions, concessions qui peuvent faire dévier ces questions des objectifs pédagogiques visés.

Le contexte de la prise de décision est sans aucun doute l'aspect le plus important à consi-

dérer dans la distinction qu'il faut établir entre la mesure normative et la mesure critériée. Dans certaines situations, la prise de décision implique le choix d'un nombre limité de candidats parmi un très grand nombre. C'est ce qui se produit en pratique lorsqu'une institution d'enseignement ne peut admettre chaque année qu'un nombre limité d'étudiants parmi un très grand nombre d'étudiants qui ont demandé leur admission. Le marché du travail comporte une situation semblable lorsqu'il s'agit de choisir un candidat parmi plusieurs qui postulent un emploi. Dans de telles situations où la prise de décision implique un choix limité, on accorde la préférence aux meilleurs candidats qui se présentent et la décision est précédée la plupart du temps d'une évaluation normative. Cette évaluation fait suite à une épreuve ou à un examen conçus dans le but de faire ressortir les différences individuelles. Il y a d'autres situations cependant où la position relative d'un individu dans un groupe donné n'est pas pertinente à une prise de décision. Par exemple, pour devenir pilote d'avion, un minimum de perception visuelle est requis. Le fait de savoir qu'un candidat possède la meilleure vision dans un groupe donné ne garantit pas qu'il possède le minimum exigé. Dans un contexte comme celui-ci, la décision d'admettre le candidat (en supposant qu'il n'y a pas de contraintes dans le nombre d'admissions possibles), doit être précédée d'une évaluation du candidat qui est faite en relation avec un seuil de perception visuelle déterminé selon les exigences du métier et non pas en relation avec la perception visuelle d'autres individus. Dans le domaine de l'apprentissage scolaire, la même situation se présente si on se préoccupe de déterminer le degré d'atteinte de certains objectifs pédagogiques. Au niveau de chaque objectif, nous pouvons être intéressés à savoir si un étudiant possède un minimum de compétence, et il n'y a pas de restrictions quant au nombre d'étudiants qui peuvent avoir ce minimum de compétence. Dans un contexte comme celui-là, une mesure conduisant à une évaluation critériée est la plus appropriée.

Nous venons de voir avec quelques exemples que dans certains cas la prise de décision peut être précédée d'une évaluation normative, alors qu'en d'autres cas la prise de décision peut être précédée d'une évaluation critériée. C'est véritablement le contexte de cette prise de décision qui nous permet de dire si une épreuve de rendement est normative ou critériée.

Évaluation formative & évaluation sommative

La notion de critère de comparaison pouvant servir de base à l'évaluation d'un étudiant, nous a amenés à établir une distinction importante entre deux types de mesure: la mesure faite en vue d'une évaluation normative et la mesure faite en vue d'une évaluation critériée. Cette distinction est utile lorsque nous nous limitons à l'évaluation d'un étudiant réalisée à un moment très précis. Si nous considérons l'ensemble des évaluations auxquelles chaque étudiant peut être soumis au cours de sa scolarité, il nous faut maintenant distinguer entre deux approches: l'approche formative et l'approche sommative. Les deux approches se distinguent l'une de l'autre selon trois points de vue: premièrement, selon la nature des décisions qui sont prises à l'endroit de chaque étudiant; deuxièmement, selon l'instant précis où l'évaluation est réalisée dans le déroulement d'un programme d'études ou d'un cours et, troisièmement, selon la nature même des habiletés qui sont évaluées.

L'évaluation formative est un système d'évaluation qui consiste à recueillir, à plusieurs occasions pendant le déroulement d'un programme d'études ou d'un cours, des informations utiles dans le but de vérifier périodiquement la qualité de l'apprentissage des étudiants. À chaque occasion, l'évaluation est conçue dans le but de décider si un étudiant peut continuer dans le programme ou dans le cours, ou s'il doit retourner en arrière pour combler des lacunes du côté de son apprentissage. Une telle décision est prise par l'enseignant et dans certains cas par l'étudiant lui-même. En évaluation formative, le contenu de chaque évaluation faisant suite à une unité d'enseignement doit être en relation directe avec des objectifs pédagogiques clairement énoncés et qui sont particuliers ou spécifiques à cette unité. La préoccupation majeure de l'évaluation formative est de mesurer, pour chaque étudiant, le degré d'atteinte de chacun des objectifs et aussi de comparer ce degré d'atteinte avec un seuil de performance déterminé. La mesure permettant une évaluation critériée de l'étudiant est sans doute le type de mesure le plus approprié, soit pour décider si l'étudiant doit entreprendre une nouvelle unité d'enseignement, soit pour lui prescrire un enseignement correctif s'il y a lieu.

L'évaluation sommative est une évaluation qui survient à la fin d'une étape importante dans la scolarité de l'étudiant comme la fin d'un cours, la fin d'un programme d'études ou d'un cycle. Contrairement à l'évaluation formative, l'évaluation sommative ne vise pas à dépister des difficultés d'apprentissage dans la perspective d'y apporter des correctifs. Elle a plutôt un caractère terminal et elle est souvent associée à des activités comme la certification, l'attestation de capacités, la promotion, le diplôme, etc. Les prises de décision reliées à l'évaluation sommative sont nombreuses et ne suivent pas nécessairement de façon immédiate le moment de l'évaluation. Par exemple, un étudiant est évalué à la fin de ses études de niveau secondaire, mais cette évaluation influencera éventuellement l'obtention d'un emploi quelques années plus tard, lorsqu'il se présentera dans le monde du travail. C'est pourquoi, dans le contexte de l'évaluation sommative, les institutions d'enseignement produisent des rapports d'évaluation qui sont transmis aux parents et aux administrateurs et qui permettent dans certains cas de différer la prise de décision. Nous serions tentés d'établir un lien nécessaire entre l'évaluation sommative et la mesure normative. Il est vrai que le contexte de certaines prises de décision qui font suite à une évaluation sommative commande habituellement une comparaison entre plusieurs étudiants, ce qui est du domaine de la mesure normative. Il n'est pas exclu cependant que le responsable d'un service d'admission ou qu'un employeur soient intéressés à connaître les capacités ou les habiletés maîtrisées par un étudiant, à la fin d'une étape importante de sa scolarité, dans le but de les comparer à celles qui sont jugées indispensables au succès dans un programme d'études ou dans une profession. Le contexte de la décision peut justifier l'emploi de mesures critériées dans l'évaluation sommative des étudiants. Enfin, le contenu de l'évaluation sommative implique des habiletés plutôt complexes et pouvant être évoquées dans une foule de situations variées. En évaluation formative, on s'intéresse à des habiletés qui sont particulières ou spécifiques à une petite unité d'enseignement et d'apprentissage. Cette dernière distinction entre l'évaluation sommative et l'évaluation formative n'est pas facile à établir dans la pratique, étant donné la complexité du domaine traitant des objectifs pédagogiques lorsqu'il s'agit de distinguer un objectif général d'un objectif spécifique.

Avant de terminer cet exposé, il nous reste deux termes à aborder qui peuvent faire l'objet d'une certaine confusion. Il s'agit de l'évaluation continue et de l'évaluation diagnostique. L'évaluation continue présente toutes les apparences de l'évaluation formative mais, avant de l'affirmer, il nous faut prendre certaines précautions. Quant à l'évaluation diagnostique, qui elle aussi semble s'apparenter à l'évaluation formative, elle est devenue, avec la création de cette dernière, une entité distincte.

À cause du caractère périodique des évaluations qui y sont réalisées, l'évaluation dite continue présente toutes les apparences extérieures de l'évaluation formative. Si on donne le nom d'évaluation continue à un système qui consiste à consigner et à accumuler plusieurs évaluations partielles dans la perspective d'une évaluation globale, on ne saurait qualifier de formative une telle évaluation. Il s'agit plutôt d'une modalité de l'évaluation sommative par laquelle l'évaluation finale est remplacée ou complétée par une somme d'évaluations partielles espacées dans le temps. Pour être considérée comme formative, l'évaluation dite continue doit être caractérisée par un ensemble de décisions à court terme qui concernent la qualité de l'apprentissage de chaque étudiant. De plus, l'accumulation d'évaluations partielles est une activité superflue en évaluation formative, étant donné que chaque évaluation est suivie immédiatement d'une prise de décision. Il demeure possible que, dans certains milieux, on veuille donner le nom d'évaluation continue à une forme d'évaluation véritablement formative. Seule une analyse des prises de décision nous permettrait de l'affirmer.

L'évaluation diagnostique se distingue de l'évaluation formative par l'objet du diagnostique et par le moment où elle est faite par rapport à un cours. Alors que l'évaluation formative a pour but de dépister les difficultés dans l'apprentissage, l'évaluation diagnostique a pour but de prévenir certaines difficultés en vérifiant la maîtrise d'habiletés pré-requises pour l'atteinte des objectifs d'un cours donné. L'objet de l'évaluation diagnostique est différent de celui de l'évaluation formative: elle porte sur les pré-requis et non sur les erreurs d'apprentissage. C'est pourquoi l'évaluation diagnostique est généralement réalisée au début d'une séquence importante d'enseignement.

M. S. - 0.

Conclusion

Dans cet article, nous avons tenté de définir et d'expliquer les principaux concepts de base qui s'inscrivent dans les approches les plus récentes à l'évaluation des étudiants. Nous avons voulu éviter la présentation sèche d'un vocabulaire technique et la spéculation sur des problèmes soulevés par l'évaluation des étudiants en milieu scolaire. Pour ce faire, nous avons abordé des modalités de mesure et d'évaluation plutôt que des techniques. De plus, nous avons exposé un point de vue fonctionnel de l'évaluation qui est basé sur la notion de prise de décision, tout en laissant le lecteur libre d'accepter ou non le bien-fondé de certaines prises de décision ●

Références

- Airasian, Peter W. et Madaus, George F., *Criterion-Referenced Testing in the Classroom*, Measurement in Education, National Council on Measurement in Education, Vol. 3, N° 4, 1972.
- Bloom, B.S., Hastings, J.T. et Madaus, G.F., *Handbook on Formative and Summative Evaluation of Student Learning*, New York: McGraw Hill Book Co., 1971.
- Popham, W. James, Editor, *Criterion-Referenced Measurement, an introduction*, Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.
- Sullivan, Howard J., « Objectives, Evaluation and Improved Learner Achievement », *Monographs Series on Curriculum Evaluation*, N° 3 (Instructional Objectives), Chicago: Rand McNally and Co., 1969, pp. 81-87.
- Vanhoven, James B., « Reporting Pupil Progress: A Broad Rationale for New Practices », *Phi Delta Kappan*, Vol. 53, N° 6, 1972, pages 365 et 366.

Les modèles d'évaluation de l'apprentissage

France Fontaine

Les termes de base tels que: évaluation formative, évaluation sommative, évaluation critériée et évaluation normative ont déjà été définis dans le texte *Évaluer? Mesurer?* Le présent texte se veut donc une suite explicitant les modèles d'évaluation dans le contexte plus restreint de l'apprentissage de l'étudiant.

L'évaluation sommative est habituellement liée à un cadre de référence normatif, tandis que l'évaluation formative est plus souvent pratiquée dans un cadre de référence critérié. On retrouvera donc, pour l'évaluation sommative et normative, puis pour l'évaluation formative et critériée, les caractéristiques, problèmes et éléments de solution afférents à ces modèles. On exposera ensuite brièvement les possibilités de modèles mixtes.

L'ÉVALUATION SOMMATIVE ET NORMATIVE

Rappelons qu'on a défini *l'évaluation sommative* comme étant un jugement définitif qui met un terme à une activité. L'évaluation a donc pour but de déterminer si un étudiant a réussi ou non, à des fins de certification, de promotion et ainsi de suite. Elle se situe à la fin de la période consacrée à l'enseignement-apprentissage, c'est-à-dire à la fin du module, du cours, du programme, alors qu'un retour en arrière n'est plus possible.

Lorsqu'on utilise *un cadre de référence normatif* pour poser un jugement, c'est-à-dire que l'on vise à situer la position d'un objet dans sa catégorie (en ce qui concerne la caractéristique envisagée), on identifie le degré de réussite d'un étudiant par des comparaisons avec ses pairs. On se réfère à des normes.

Caractéristiques

a) L'évaluation sommative ayant un caractère définitif, c'est-à-dire une grande importance pour l'avenir de l'étudiant, elle présente les caractéristiques suivantes :

- elle est *valide* c'est-à-dire qu'elle doit mesurer tout ce qu'il faut mesurer et cela seulement;
- elle est *fidèle*, c'est-à-dire que les mesures doivent être précises et stables;
- elle se doit d'être *objective* c'est-à-dire d'assurer un traitement égal pour tous;
- elle est nécessairement *formelle*, c'est-à-dire que les règles d'évaluation doivent être clairement établies et observées;
- elle ne porte habituellement que sur un *échantillon des objectifs* du cours, c'est-à-dire sur les objectifs jugés les plus importants;
- elle n'a que *peu de conséquences sur la planification* du cours.

b) Comme l'évaluation normative vise à situer les étudiants les uns par rapport aux autres selon leur apprentissage, il faut choisir des instruments de mesure qui facilitent les comparaisons entre les étudiants. Ainsi, pour un test écrit, on choisira les questions ou items qui discriminent le plus, qui distinguent le plus les étudiants les uns des autres. C'est pourquoi on éliminera d'un test les questions dites faciles, parce que répondues par presque tous les étudiants, ou celles jugées trop difficiles, parce que répondues par un très petit nombre d'étudiants.

Problèmes

1° Souvent, les instruments de mesure (examens, tests, etc.) utilisés ne sont guère valides, fidèles et objectifs.

2° On analyse rarement les instruments de mesure à partir des réponses des étudiants, par manque de temps, de moyens ou de motivation.

3° Les points sur lesquels porte l'évaluation constituent parfois un échantillon plus ou moins représentatif de l'ensemble des objectifs du cours: on n'évalue pas les objectifs essentiels.

4° Il arrive qu'on accorde une importance absolue à des instruments de mesure, qu'ils soient bien construits ou non. Ceci mène à l'automatisme dans la prise de décision: par exemple, lorsque l'étudiant obtient 60% et plus à son test, il réussit, sinon il échoue.

5° Des comportements peu souhaitables sont fréquemment liés à l'évaluation sommative. L'étudiant peut travailler en fonction de sa note finale; il peut chercher à deviner ce qui est important pour le professeur si celui-ci ne l'a pas suffisamment explicité; il arrive souvent qu'il y ait bourrage de crâne à la veille de l'examen.

6° Une trop grande anxiété de la part de certains étudiants en période d'examens diminue la fiabilité des résultats.

7° Comme on identifie les difficultés d'apprentissage des étudiants qu'à la fin de la période d'enseignement-apprentissage, il est trop tard pour apporter les correctifs qui s'imposent. Il peut donc arriver qu'un étudiant poursuive ses études malgré certaines lacunes ou qu'on l'oblige à recommencer tout son cours même si ses lacunes sont concentrées dans telle ou telle partie du cours.

8° L'ambiguïté est parfois grande entre le rôle du professeur comme facilitateur de l'apprentissage et son rôle de juge de ces apprentissages.

9° L'étudiant, n'ayant été contrôlé qu'à la fin de l'apprentissage, n'a guère pu développer de critères pour s'auto-évaluer¹ correctement. Une des conséquences en est que plusieurs étudiants considèrent injustifiés leurs résultats scolaires.

10° Lorsqu'on évalue suivant le cadre de référence normatif, l'échec ou le degré de réussite d'un étudiant dépend de la distribution des résultats de l'ensemble des étudiants. Ainsi, selon que le groupe auquel on le compare est plus ou moins fort, il pourra, à la limite, échouer ou réussir.

Éléments de solution

Soulignons d'abord que si certains problèmes peuvent être résolus, partiellement ou complètement, certains autres sont inhérents à cette forme d'évaluation (par exemple, les problèmes 7 et 9). Il faut s'attendre à ce que ces problèmes subsistent si l'évaluation n'a pas pour but de les résoudre.

- Afin d'améliorer la construction des instruments de mesure, on pourrait mieux identifier les objectifs du cours, les critères de réussite et distinguer entre les objectifs terminaux² et les objectifs intermédiaires. Comme l'évaluation sommative se limite habituellement à une partie des objectifs, il est possible de ne la faire porter que sur les objectifs jugés importants, par exemple les objectifs terminaux.

.....

- Il est de plus en plus facile d'analyser les instruments de mesure (surtout composés de questions à choix multiple, c'est-à-dire de questions offrant un choix de réponses) car le nombre de programmes de traitement informatisés créés à cette fin est sans cesse grandissant. On peut ainsi accumuler les statistiques concernant les questions, éliminer ou reformuler les questions mal construites et le reste.
- Il faut s'assurer qu'il y a accord sur la réponse à donner à une question: trop souvent il y a divergence chez les professeurs. C'est l'étudiant qui en paye les frais.

- On peut augmenter l'objectivité dans la correction de tests ou de travaux en préparant des instruments tels qu'une liste des bonnes réponses avec le nombre de points à accorder pour chacune, une liste de critères d'évaluation détaillés et pondérés et le reste.
- Lorsqu'on assure aux étudiants de bonnes conditions d'administration des instruments de mesure (instructions claires, environnement favorable, etc.), on contribue à augmenter la fidélité de ces instruments. Dans le cas de questions à choix multiple, on peut diminuer les erreurs de mesure dues au choix d'une réponse au hasard en accordant suffisamment de temps aux étudiants pour répondre, en évitant les questions trop difficiles, en posant d'abord les questions faciles puis les questions difficiles et ainsi de suite.
- Comme tout instrument d'évaluation de l'apprentissage ne mesure pas directement l'objet à évaluer, il y a toujours erreur de mesure. Donc le résultat observé d'un étudiant n'est pas nécessairement son résultat véritable. On peut cependant estimer cette erreur de mesure à partir du coefficient de fidélité du test. On peut ainsi calculer, avec un certain degré de probabilité, un intervalle de confiance autour du résultat observé, dans lequel se situerait le résultat véritable de l'étudiant. On peut, de la sorte, diminuer l'arbitraire de la note de passage en appliquant ce calcul d'intervalle de confiance à la note de passage.
- Si le professeur communique à l'étudiant ses intentions pédagogiques par un énoncé explicite des objectifs du cours, en indiquant si certains objectifs sont prioritaires et si les instruments de mesure sont conformes aux objectifs, l'étudiant n'aura pas à deviner ce qui est important pour le professeur. Il pourra se concentrer sur les apprentissages à faire et il est probable qu'il sera moins anxieux par rapport à l'évaluation.
- Il est parfois possible d'éviter le bourrage de crâne à la veille des examens en vérifiant si les objectifs que l'on vise sont ou non des objectifs de connaissance. Lorsqu'il ne s'agit pas de tels objectifs, on peut en informer les étudiants et construire les instruments de mesure en conséquence. Il est peut-être possible aussi de diminuer la tendance au bachotage par un plus grand nombre d'évaluations sommatives, c'est-à-dire par l'évaluation continue⁴.
- Comme il n'est pas obligatoire que l'évaluation sommative soit également normative, on peut se demander si le cadre de référence normatif est bien celui qui convient, si on en tire bien les renseignements désirés. Serait-il plus avantageux de situer l'apprentissage de l'étudiant par rapport aux objectifs ou est-il plus utile de comparer l'étudiant aux autres étudiants de son groupe?

L'ÉVALUATION FORMATIVE ET CRITÉRIÉE

L'évaluation formative vise à améliorer l'apprentissage: elle a lieu pendant l'apprentissage et permet un retour en arrière si des corrections s'imposent. Ce qui la distingue de l'évaluation sommative, c'est qu'une intervention est encore possible; ce n'est pas un jugement définitif. Elle est en elle-même une activité d'apprentissage, l'étudiant apprenant par cette évaluation, qui le renseigne sur ses succès et échecs.

Lorsqu'un *cadre de référence critérié* est utilisé pour poser un jugement, on cherche à déterminer si l'objet est satisfaisant par rapport à un critère, indépendamment de sa position dans sa classe.

On se demande si l'étudiant atteint les critères prédéterminés, et ce, sans référence aux autres étudiants de son groupe. On se réfère donc à des critères plutôt qu'à des normes.

Caractéristiques

a) Comme l'évaluation formative a pour but d'aider l'étudiant dans son apprentissage, elle présente les caractéristiques et entraîne les conséquences suivantes:

- Elle sera suffisamment fréquente pour que l'étudiant ait le temps de revenir sur des apprentissages, s'il y a lieu. Ceci lui permettra également de ne pas s'égarer trop longtemps sur une fausse piste, comme l'étudiant qui ne comprendrait pas les instructions données par le professeur pour un travail ou qui procéderait de façon erronée pour résoudre un certain type de problèmes. Qu'il s'agisse de renforcement ou de corrections à apporter par d'autres activités d'apprentissage, il faut tarder le moins possible à informer l'étudiant de ses succès et de ses erreurs. Il ne faut cependant pas confondre cette évaluation avec une forme d'évaluation qu'on peut qualifier de cumulative sommative, souvent appelée continue⁵, qui ne permet pas de revenir sur des apprentissages.
- Elle exige que les résultats de l'évaluation soient connus rapidement des étudiants, pour les mêmes raisons que celles énoncées au sujet de la fréquence de l'évaluation. Il n'est pas rare de voir les résultats d'une évaluation sommative remis plusieurs jours, voire plusieurs semaines, après les épreuves d'évaluation. Une telle pratique n'aide guère l'étudiant à combler les lacunes de ses apprentissages, surtout dans le cas où il est limité à une période déterminée (souvent quinze semaines) pour atteindre les objectifs du cours.
- Elle nécessite la prescription d'activités d'apprentissage si l'objectif n'est pas atteint. Il peut s'agir de reprendre certaines des mêmes activités d'apprentissage ou encore de procéder à de nouvelles activités d'apprentissage mieux adaptées aux besoins de l'étudiant (par exemple: visionner un film au lieu des lectures prescrites). Le professeur doit penser à ces autres activités d'apprentissage et, s'il y a lieu, concevoir et réaliser le matériel qu'elles supposent.

- Elle est transparente, c'est-à-dire que les étudiants doivent savoir ce qu'on attend d'eux. Il n'y a pas de secret quant à l'objet de l'évaluation.
- Elle n'est pas nécessairement formelle. Par exemple, elle peut être basée sur des instruments d'auto-évaluation⁶ mis à la disposition des étudiants avec prescription d'activités d'apprentissage. Ou encore le professeur peut aider quelques étudiants dont il a observé les difficultés d'apprentissage et révéifier ensuite les acquis.
- Elle exige assez souvent une planification de cours plus poussée que l'évaluation sommative. C'est habituellement le cas lorsque les étudiants sont nombreux ou lorsqu'on veut bien intégrer l'évaluation formative à l'enseignement, par exemple en rendant plus formelle l'évaluation formative. Il faut choisir à quel moment auront lieu les contrôles, prévoir quand et comment seront révéifiés les apprentissages des étudiants qui ont manifesté des lacunes au premier contrôle et le reste.
- Comme l'incidence de l'évaluation formative sur l'étudiant n'a pas de caractère définitif, contrairement à l'évaluation sommative, plusieurs auteurs sont d'avis qu'il n'est pas nécessaire que les instruments d'évaluation soient aussi valides et fidèles que ceux qui sont requis par l'évaluation sommative.
- Afin de déplaister les problèmes d'apprentissage, elle demande que tout objectif soit vérifié, qu'il s'agisse d'un objectif intermédiaire, d'un objectif jugé facile à atteindre, etc. Le rôle de l'évaluation formative peut être des plus importants lorsqu'il y a des séquences d'apprentissage.
- Elle conduit assez souvent à l'individualisation de l'enseignement⁷, au respect du rythme d'apprentissage. On permet à l'étudiant de reprendre des apprentissages sans le contraindre à le faire en dehors des heures de cours, sans le soumettre au rythme du groupe. Dans ce cas, le rôle du professeur est celui d'un facilitateur de l'apprentissage plutôt que d'un dispensateur de connaissances.

b) Comme l'évaluation critériée vise non pas à situer l'étudiant par rapport aux autres étudiants, mais à dire si l'étudiant atteint ou non les critères prédéterminés :

- on n'éliminera pas des instruments les questions trop faciles ou trop difficiles; il suffit qu'elles renseignent sur des apprentissages pertinents;
- elle se fera habituellement en fonction d'objectifs spécifiques⁸ et on dira à l'étudiant qu'il a réussi ou non à atteindre tel objectif.

.....

Problèmes

Les deux premiers problèmes mentionnés pour l'évaluation sommative et normative (portant sur la validité, la fidélité, l'objectivité et l'analyse des instruments), peuvent aussi se présenter ici, bien que, comme on l'a mentionné dans les caractéristiques de l'évaluation formative, ces problèmes ont peut-être moins d'acuité. Fréquemment, certains des moyens envisagés pour résoudre ces deux problèmes font déjà partie de la démarche d'évaluation formative et critériée, par exemple rédiger des objectifs précis pour construire des instruments qui correspondent à ceux-ci.

De plus, on rencontre les difficultés suivantes :

1° La *tâche du professeur* est assez souvent accrue. (Parfois il doit consacrer plus de temps à la planification du cours, à la construction d'instruments d'évaluation, à la conception d'un système de correction ou, lors du déroulement du cours, au choix d'activités d'apprentissage différentes, à la correction, etc.)

2° L'*attitude du professeur* est un des facteurs qui déterminent le succès ou l'échec d'un processus d'évaluation formative dans lequel le professeur joue un rôle actif. L'étudiant doit sentir que son professeur désire l'aider à atteindre les objectifs du cours. Mais, tous les professeurs ne sont pas également à l'aise dans ce rôle.

3° Les étudiants éprouvent aussi des difficultés à modifier leurs *attitudes par rapport à l'évaluation et au professeur*. Par exemple, plusieurs croient qu'il leur est avantageux de cacher au professeur leurs difficultés d'apprentissage. (Lorsque l'évaluation formative est jumelée avec le «mastery learning¹⁰» et l'individualisation de l'enseignement¹¹, il n'est pas rare qu'on ait à faire face à des étudiants insécures qui vivent peut-être ce système pour la première fois).

4° Comme la *durée permise pour les apprentissages* d'un cours est généralement prédéterminée (habituellement, on dispose d'une quinzaine de semaines), on limite l'étudiant dans ses possibilités de faire des retours sur ses apprentissages. Autrement dit, le plein épanouissement de l'évaluation formative n'est pas facile à réaliser dans notre système éducatif.

5° Un minimum de *planification* est requis lorsque l'évaluation formative est intégrée à un cours. Ce minimum augmente au fur et à mesure qu'on rend plus formelle l'évaluation formative, qu'on s'adresse à un plus grand nombre d'étudiants.

Par exemple, il faut alors assurer une certaine souplesse dans le cours afin de permettre à l'étudiant de faire les réajustements nécessaires. Il faut choisir la fréquence des contrôles d'évaluation formative de telle sorte que l'étudiant reçoive fréquemment de l'information sur ses apprentissages sans pour cela que l'évaluation soit le seul moyen d'apprentissage, ou presque. L'équilibre n'est pas toujours facile à trouver. Il faut aussi préciser les modalités de recontrôle des étudiants qui ont manifesté des lacunes.

Il faut prendre ces décisions de sorte que le tout soit cohérent. Le «design» d'un tel cours peut être difficile à réaliser: le choix de mauvaises modalités d'application peuvent rendre nuls les avantages de l'évaluation formative.

6° Il n'est pas nécessairement facile d'assurer des *moyens d'évaluation équivalents* pour évaluer les étudiants ayant manifesté des lacunes dans leurs apprentissages, lors d'un contrôle précédent.

Éléments de solution

Comme pour l'évaluation sommative, il ne s'agit ici que de solutions partielles à des problèmes parfois complexes.

- Lorsque plusieurs professeurs peuvent travailler à un même cours, il est plus facile de mettre sur pied un système d'évaluation formative. Ils peuvent même assurer le cours en équipe. Dans le cas d'un professeur travaillant seul, celui-ci pourra planifier et expérimenter une partie du cours seulement, ou intégrer peu à peu les éléments d'un système d'évaluation formative de plus en plus complet tout en s'assurant de la cohérence à chacune des étapes. Ceci lui permettra d'identifier et de résoudre les problèmes au fur et à mesure.

Un professeur peut aussi chercher à alléger sa tâche de correction par des auto-tests, par la mise sur pied d'un système de correction automatisée, par l'utilisation de questions à réponses brèves, etc., pour autant que ces moyens soient compatibles avec les objectifs du cours.

- Le professeur devrait se demander, même s'il est convaincu des avantages de l'évaluation formative, s'il accepte réellement le rôle d'aide à l'étudiant et si ses comportements face à celui-ci lui permettent de jouer efficacement ce rôle. Ainsi, un professeur qui semble manifester un certain mépris pour ses étudiants ne pourra pas vraiment réussir dans ce mode d'évaluation même s'il s'agit d'une fausse perception par les étudiants. Il lui faudra d'abord modifier les comportements qui sous-tendent cette perception étudiante (par exemple, éviter de ridiculiser ou de rabrouer un étudiant qui éprouve des difficultés d'apprentissage).
- Le professeur doit expliquer clairement aux étudiants les buts de l'évaluation formative et les inciter, par son attitude, par sa disponibilité, par les modalités du cours, à parfaire leurs apprentissages. Il doit faire croître, chez les étudiants, le sentiment de responsabilité face à leurs apprentissages et au contrôle de ceux-ci. Mais ce n'est pas facile. (Lorsqu'on a jumelé l'évaluation formative avec «le mastery learning¹²» et l'individualisation de l'enseignement¹³, on peut rassurer l'étudiant en lui faisant sentir qu'il est suffisamment encadré, en lui exposant les résultats de cours similaires, etc.)
- Fréquemment, l'administration n'accorde que les quinze semaines du semestre et la période des examens pour compléter les dossiers étudiants. Il est parfois possible d'inscrire la mention «incomplet» ou «remise», accordant ainsi un délai à l'étudiant. Il peut alors terminer les apprentissages prévus entre les deux sessions ou à la session suivante, sans être obligé de recommencer les apprentissages correspondants aux objectifs déjà atteints.

- Avant d'annoncer aux étudiants un système d'évaluation formative, le professeur doit avoir estimé les ressources disponibles, prévu le déroulement du cours, préparé son matériel (par exemple: objectifs, matériel d'apprentissage, instruments d'évaluation, dossier étudiant), précisé les modalités d'application et le reste. C'est donc dire que le professeur doit s'assurer d'une certaine disponibilité auprès des étudiants durant le cours.
- La réalisation de banques de questions ou d'items¹⁴ et les techniques d'analyse d'items peuvent être des solutions à l'élaboration de moyens d'évaluation équivalents. (Les banques d'items permettent d'accumuler un grand nombre de questions codifiées en fonction des objectifs qu'elles mesurent.)

LES MODÈLES MIXTES

Puisque nos systèmes d'éducation exigent habituellement une note différenciée afin de sélectionner les meilleurs étudiants et qu'il est par ailleurs difficile, dans le contexte d'un semestre, de maintenir un système d'évaluation formative et un système d'évaluation sommative, vu le temps requis pour la préparation et l'application des instruments de mesure, on essaye de faire des compromis entre les deux systèmes. Il faut cependant préserver la cohérence. Ceci n'est pas facile, car il s'agit de deux modèles d'évaluation visant des buts différents. Il n'y a sans doute pas de solution idéale: il s'agit de choisir celle qui convient le mieux au cours et au programme de formation.

Par exemple, la réussite à un test ou travail contribue pour tel pourcentage à la note, le test ou travail pouvant être repris sans pénalité jusqu'à ce que la note atteigne un standard prédéterminé ou que l'étudiant soit satisfait de ses résultats. D'une autre façon, on peut accorder une note (comme 60% ou 70%) pour la réussite des tests ou travaux formatifs et compléter cette note par celle qui est attribuée pour un examen ou travail supplémentaire permettant de différencier les étudiants. Ce ne sont que deux exemples parmi plusieurs possibles.

Conclusion

L'évaluation sommative et l'évaluation formative de l'apprentissage répondent à des fins différentes; elles servent d'appui à des décisions distinctes. Il ne s'agit donc pas de vanter les mérites d'un système d'évaluation par rapport à l'autre, mais de choisir un système d'évaluation répondant aux besoins identifiés.

Comme l'évaluation formative, surtout lorsqu'elle est structurée, n'est pas très répandue dans notre système d'éducation, il reste à faire plusieurs recherches et expérimentations en ce domaine. Ce texte ne peut donc pas être un tableau définitif des modèles d'évaluation.

Il est à noter que les modèles d'évaluation peuvent s'appliquer à d'autres objets que l'apprentissage de l'étudiant. On peut songer, dans le domaine de l'éducation, à l'évaluation de l'enseignement donné par un professeur¹⁵ ou à l'évaluation d'un programme.

Par exemple, l'enseignement d'un professeur peut être évalué afin de décider s'il sera réengagé ou promu (évaluation sommative) ou encore afin de lui permettre d'améliorer son enseignement (évaluation formative). Ceci peut se faire en fonction de critères (évaluation critériée) ou en fonction des résultats de ses collègues (évaluation normative). L'évaluation de programme peut avoir pour but de choisir le ou les programmes qui pourront continuer à bénéficier de subsides, qui devront disparaître, etc. (évaluation sommative) ou encore déterminer les éléments du programme qui sont à réviser (évaluation formative). L'évaluation peut se faire en fonction de critères ou en fonction d'autres programmes.

Notes et références bibliographiques

1. Fontaine, France, «Il ne faut pas confondre l'autonotation et l'auto-évaluation» dans *Dossier sur l'évaluation*, pp. 29-32.
2. Fontaine, France, *Les objectifs d'apprentissage*. Montréal, Service pédagogique, Université de Montréal, 1980, 110 p..
3. Bloom, Benjamin S. et al., *Taxonomie des objectifs pédagogiques, Tome I, domaine cognitif*. Traduit de l'américain par: Lavallée, Marcel. *Taxonomy of Educational Objectives*, 1958. Montréal, Education nouvelle, 1969, 232 p..
4. Fontaine, France, «Évaluer? Mesurer?» dans *Dossier sur l'évaluation*, pp. 5-9.
5. Id., *ibid.*
6. Fontaine, France, «Il ne faut pas confondre l'autonotation et l'auto-évaluation» dans *Dossier sur l'évaluation*, pp. 29-32.
7. Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques*. Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.
8. Fontaine, France, *Les objectifs d'apprentissage*. Montréal, Service pédagogique, Université de Montréal, 1980, 110 p..
9. Felix-Larivière, Claire, Fontaine, France, «Assurer à tous la compétence» dans *Dossier sur l'évaluation*, pp. 11-15.

10. Id., *ibid.*
11. Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques.* Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.
12. Felix-Larivière, Claire, Fontaine, France, «Assurer à tous la compétence» dans *Dossier sur l'évaluation.* pp. 11-15.
13. Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques.* Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.
14. Bernard, Huguette, Normand, Serge, «La banque d'objectifs et d'items: un exemple d'application» dans *Dossier sur l'évaluation,* pp. 123-130.
15. Fontaine, France, «Démarche pour évaluer l'enseignement» dans *Dossier sur l'évaluation,* pp. 51-71.

ACTIVITE D'EVALUATION FORMATIVE I

SECTION I

Pour chacun des exemples suivants, écrivez sur la feuille-réponses la lettre "M" s'il s'agit du résultat d'une mesure, et la lettre "E" s'il s'agit du résultat d'une évaluation.

- 1- Un métal de grande densité.
- 2- Un élève médiocre.
- 3- Une note de 20 points à un examen de 20 questions.
- 4- Un élève habile.
- 5- Un objet pesant 0.2 milligramme.

SECTION II

Voici des contextes variés d'utilisation d'instruments de mesure. Identifiez de quel type d'évaluation il s'agit dans chacun des contextes en associant sur la feuille-réponses, le numéro de l'énoncé à la lettre représentant le type d'évaluation utilisé.

CONTEXTES

TYPE D'EVALUATION

- | | |
|---|----------------------------|
| 6- Une épreuve de rendement sur le balancement d'équations en chimie qui permet de recommander à chaque étudiant la reprise de certains exercices selon une difficulté typique éprouvée par cet étudiant. | a) Evaluation formative |
| 7- Un test apparaît au tout début d'un livre d'auto-enseignement et a pour but de conseiller le lecteur sur le ou les chapitres que celui-ci devrait aborder en premier lieu et aussi sur le ou les chapitres qu'il peut s'abstenir de lire (compte tenu de connaissances acquises antérieurement). | b) Evaluation sommative |
| | c) Evaluation diagnostique |

CONTEXTES (suite)

TYPE D'EVALUATION (suite)

- | | |
|---|----------------------------|
| 8- Une épreuve de rendement couvrant l'ensemble des objectifs poursuivis lors d'une étape, administrée à la fin de cette étape, et dont les résultats seront convertis en rangs stanines pour être ensuite consignés dans un bulletin cumulatif. | a) Evaluation formative |
| 9- Un professeur questionne oralement les étudiants de sa classe afin de décider s'il peut poursuivre sa matière ou s'il doit revenir sur des notions qu'il juge non suffisamment maîtrisées. | b) Evaluation sommative |
| 10. Un test de rendement (dans une matière donnée) a été conçu de façon à repérer les difficultés que les élèves n'ont pu surmonter l'année précédente afin d'organiser une révision qui assure la maîtrise des pré-requis avant d'entreprendre le cours. | c) Evaluation diagnostique |

SECTION III

Complétez les phrases suivantes en inscrivant sur la feuille-réponses le mot adéquat.

- 11- L'évaluation faite dans le but d'améliorer l'apprentissage et l'enseignement s'appelle l'évaluation-----.
- 12- On appelle-----une évaluation sommative cumulative.
- 13- L'évaluation qui vise à situer la performance d'un élève dans un groupe en fonction d'une habileté particulière s'appelle l'évaluation-----.
- 14- La-----consiste en l'application d'une règle d'attribution de symboles à des objets.
- 15- L'évaluation qui vise à déterminer si un élève a réussi ou non, à des fins de promotion ou de certification s'appelle évaluation-----.
- 16- Porter un jugement de valeur, c'est-----.

SECTION III (suite)

- 17- Lorsque l'évaluation vise à déterminer si l'apprentissage d'un élève est satisfaisant par rapport à un seuil de maîtrise et non par rapport au reste de la classe, il s'agit d'une évaluation-----.
- 18- L'évaluation-----se fait avant la période consacrée à l'enseignement et a pour but de déterminer quels seront les prérequis qu'il faudra revoir.

SECTION IV

Indiquez en cochant les cases appropriées sur la feuille-réponses si les énoncés suivants sont vrais ou faux.

- 19- Je peux évaluer sans mesurer mais je me base alors sur mes préjugés ou mes opinions.
- 20- L'évaluation sommative est ponctuelle, l'évaluation formative est continue.
- 21- L'évaluation est un moyen au service d'une prise de décision.
- 22- L'unité de mesure en éducation, soit le point, est une unité de mesure qui peut servir de standard.
- 23- L'évaluation diagnostique se fait aux mêmes moments que l'évaluation formative.

SECTION V

Choisissez la bonne réponse et inscrivez sur la feuille-réponses la lettre correspondante.

- 24- Lorsque la décision qu'on doit prendre implique le choix d'un nombre limité de candidats, le type d'évaluation utilisé sera la plupart du temps
- a) normatif
 - b) formatif
 - c) critérié
 - d) diagnostique
 - e) sommatif

SECTION V (suite)

25- Lorsque la décision qu'on doit prendre implique la présence d'un minimum de compétence, le type d'évaluation le plus approprié est l'évaluation.

- a) sommative
- b) normative
- c) formative
- d) critériée
- e) diagnostique

26- Le système d'évaluation qui consiste à recueillir pendant le déroulement d'un cours des informations utiles dans le but d'attester de la qualité de l'apprentissage s'appelle

- a) évaluation sommative
- b) évaluation continue
- c) évaluation critériée
- d) évaluation formative
- e) évaluation diagnostique

SECTION VI

Pour chacune des situations suivantes inscrivez sur la feuille-réponses:

- a) si la situation exige ou implique une évaluation (ou mesure) critériée
- b) si la situation exige ou implique une évaluation (ou mesure) normative
- c) si les données sont insuffisantes pour dire que la situation implique une évaluation critériée ou normative.

27- SITUATION A:

Dans une polyvalente on a mis sur pied deux types de classes de niveau Secondaire I: (a) des classes régulières pour accueillir les élèves maîtrisant un certain nombre d'objectifs considérés comme prérequis à la poursuite d'études secondaires et (b) des classes spéciales dites de récupération pour les élèves ne maîtrisant pas ces objectifs. Les élèves qui proviennent du niveau élémentaire subissent une épreuve de rendement

SECTION VI - SITUATION A (suite)

et doivent obtenir un résultat minimum pour être admis dans une classe régulière. Le nombre de classes de chaque catégorie n'est pas déterminé à l'avance mais dépend du nombre d'étudiants ayant atteint ou non le résultat fixé comme minimum.

28- SITUATION B:

Un organisme dispense un programme de secourisme et tient compte des informations suivantes pour l'admission des candidats: l'âge, le nombre d'années de résidence au Canada, la citoyenneté, le sexe et la note brute à un examen de connaissances générales en matière d'accidents.

29- SITUATION C:

D'après un règlement du Ministère de l'éducation, les commissions scolaires doivent admettre dans les classes de l'élémentaire tous les enfants qui ont atteint un âge donné au 1er septembre.

30- SITUATION D:

Une personne connaît ses performances individuelles à plusieurs épreuves de course à pied (100 mètres, 500 mètres, 100 mètres-haie etc.). Parmi ces types d'épreuves elle veut connaître celui dans lequel ses chances de succès sont les meilleures si elle devait participer à des épreuves de compétition sportive dans le domaine de la course à pied.

31- SITUATION E:

Une association regroupant des spécialistes en météorologie décide de combler un poste laissé vacant par le départ d'un membre. Il y a quinze candidats qui se présentent et un seul sera accepté.

32- SITUATION F:

Dans une institution d'enseignement l'évaluation finale des étudiants pour chaque cours est traduite par les lettres A,B,C,D et E. La politique de l'institution stipule que dans chaque cours 10% des étudiants doivent recevoir la lettre A, 20% la lettre B, 40% la lettre C, 20% la lettre D et 10% la lettre E.

(votre nom s.v.p.)

ACTIVITE D'EVALUATION FORMATIVE I

Feuille-réponses

Section I

1. () 2. () 3. () 4. () 5. ()

Section II

6. () 7. () 8. () 9. () 10. ()

Section III

11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____

Section IV

19. Vrai () Faux () 22. Vrai () Faux ()
20. Vrai () Faux () 23. Vrai () Faux ()
21. Vrai () Faux ()

Section V.

24. () 25. () 26. ()

Section VI

27. () 28. () 29. () 30. ()
31. () 32. ()

ACTIVITE D'EVALUATION FORMATIVE I

Corrigé

Section I

1. (E) 2. (E) 3. (M) 4. (E) 5. (M)

Section II

6. (a) 7. (c) 8. (b) 9. (a) 10. (c)

Section III

- 11. formative
- 12. continue
- 13. normative
- 14. mesure
- 15. sommative
- 16. évaluer
- 17. critériée
- 18. diagnostique

Section IV

19. Vrai (X) Faux () 22. Vrai () Faux (X)
20. Vrai () Faux (X) 23. Vrai (.) Faux (X)
21. Vrai (X) Faux ()

Section V.

24. (a) 25. (d) 26. (a)

Section VI

27. (a) 28. (a) 29. (a) 30. (b)
31. (b) 32. (b)

EVALUATION FORMATIVE I

Grille d'interprétation du résultat au test

A. Mesure et évaluation

1 point pour chaque bonne réponse aux questions: 1-2-3-4-5-14-16-19-21-22.

Mon résultat: /10

Seuil de maîtrise: 9/10

B. Evaluation formative et sommative

1 point pour chaque bonne réponse aux questions: 6-7-8-9-10-11-15-20-23-26.

Mon résultat: /10

Seuil de maîtrise 9/10

C. Evaluation normative ou critériée

1 point pour chaque bonne réponse aux questions: 13-17-24-25-27-28-29-30-31-32.

Mon résultat: /10

Seuil de maîtrise 9/10

D. Evaluation diagnostique

1 point pour chaque bonne réponse aux questions: 6-7-8-9-10-18-23.

Mon résultat: /7

Seuil de maîtrise 6/7

E. Evaluation ponctuelle et continue

1 point pour chaque bonne réponse aux questions: 12-20.

Mon résultat /2

Seuil de maîtrise 2/2

EXAMINATION FORM

Name of Candidate

Roll Number

Subject

Non-Resident
Seat of Candidate

Name of Candidate

Roll Number

Non-Resident
Seat of Candidate

Name of Candidate

Roll Number

Non-Resident
Seat of Candidate

Name of Candidate

Roll Number

Non-Resident
Seat of Candidate

Name of Candidate

Roll Number

Non-Resident
Seat of Candidate

Fiche 8732
Des. 305