

## Actes du 14<sup>e</sup> colloque de l'AQPC

# ÉVALUATION ! ÉVOLUTION ?



Où s'en va le collégial ?

7A21

Intégrer l'évaluation de l'apprentissage  
à son enseignement en 9 étapes

par  
RACINE, Serge  
professeur  
Université de Montréal



Association québécoise  
de pédagogie collégiale

Serge Racine  
 Professeur  
 Université de Montréal

## Intégrer l'évaluation de l'apprentissage à son enseignement en 9 étapes<sup>1</sup>

### Introduction

L'apprentissage est un changement dans les fonctions de l'étudiant tel que visé par les objectifs, amplifié par l'enseignement et constaté par l'évaluation, d'où l'importance de la congruence "objectifs-enseignement-évaluation". Ces phases comportent 9 étapes: (1) Partir de l'étudiant, (2) Opérationnaliser les objectifs, (3) Organiser son enseignement pour favoriser le succès, (4) Choisir un modèle d'enseignement, (5) Choisir son mode d'évaluation, (6) Élaborer le domaine, (7) Élaborer les instruments de mesure, (8) Administrer ces instruments et (9) les Réviser.

### 1. Partir de l'étudiant

Imaginons une théorie de l'éducation qui prétend que la personne apprend en agissant et que ces actions la spécialisent. La personne se forge des fonctions conformément au mécanisme de son système nerveux. Les fonctions, imaginative, affective, physique, cognitive et sociale (IAPCS) existent comme des spécialités de la personne. Telle personne devient imaginative à force d'utiliser son imagination, comme on devient droitier ou gaucher. La personne améliore ses fonctions de manière à obtenir plus d'harmonie. L'harmonie chez la personne produit une synergie, une puissance plus grande que la somme des fonctions. Cette énergie s'exprime par une plus grande capacité à résoudre des problèmes, plus d'amour, plus de santé, plus de connaissances et plus de succès avec les autres. (« La personne synergique », Racine, 1994).

Le point de départ d'une théorie est important. Nous proposons de **partir de l'étudiant**. Notre conception de la personne qui apprend détermine la manière d'envisager l'apprentissage, l'enseignement et l'évaluation. *Nous proposons une vision dynamique et holistique du développement de la personne*. Chaque être humain reçoit un bagage spécifique à travers son code génétique. Il reçoit cependant dans ce bagage un système nerveux **programmable**. La personne naît avec des milliards de neurones **disponibles**. Ces neurones se connectent entre eux, à l'occasion de ses **expériences**, formant un chemin qui s'exprime par une réponse comportementale. Chaque nouvelle activité engendre une nouvelle association de dendrites ou un nouveau chemin neuronal (Changeux, 1983). Ainsi, nous bâtissons nos fonctions à partir de nos expériences un peu comme une spécialisation de notre système nerveux, dans le même sens que nous devenons droitiers pour écrire. **Nous devenons ce que nous faisons**. Ensuite, nous adaptons nos fonctions au besoin. C'est ce que suggère aussi Bloom (1979) quand il prétend que les expériences cognitives et affectives précédentes déterminent 65% du succès à l'activité académique suivante. Un autre 25% est déterminé par ce que fait l'étudiant durant son apprentissage. Ainsi, toute action spécialise le système nerveux de la personne et les diverses formes d'activités humaines engendrent des fonctions. Nous appelons ces fonctions "imaginative, affective, physique, cognitive et sociale". Quand nous proposons un objectif à un étudiant, nous nous adressons à toutes ces fonctions. L'apprentissage est le changement dans ces fonctions.

Du point de vue physiologique, la synergie est l'association de plusieurs organes pour l'accomplissement d'une fonction. En général, il s'agit de la mise en commun de plusieurs actions concourant à un effet unique avec une économie de moyens et d'énergie. Les cinq grandes fonctions dont se dote la personne travaillent ensemble. Imaginons un carrosse tiré par cinq chevaux. Chacun a sa propre force, par exemple 1 newton. Si un seul d'entre eux est fatigué ou malade, les autres ne peuvent donner leur plein rendement. L'énergie totale de cet attelage est une force moindre que quatre newtons. Par contre, quand les cinq travaillent ensemble, c'est-à-dire bien synchronisés, complémentaires, en harmonie et au même rythme, on s'aperçoit que chacun travaille moins de 1 newton alors que l'énergie totale dépasse les 8 newtons. C'est le même phénomène qui se produit quand une équipe de hockey travaille en harmonie. L'effet est fulgurant. C'est ce qui s'appelle la synergie; **l'ensemble donne plus d'énergie que la somme des parties**. De même une personne dont les cinq fonctions travaillent en harmonie retrouve une puissance bien plus grande que la somme des parties. C'est pourquoi ce ne sont pas les personnes les plus fortes physiquement qui sont les plus puissantes personnellement. Ni celles qui ont développé leur fonction cognitive au

<sup>1</sup> Texte de l'atelier présenté au 14e colloque de l'Association québécoise de pédagogie collégiale, le 7 juin 1994.

détriment des autres fonctions. Mais qu'est la puissance humaine? Une personne puissante est essentiellement une personne à la fois habile à créer, en amour, en santé, autonome et prospère. C'est le but de chaque être humain. L'éducation devrait favoriser l'atteinte de cet objectif parce qu'il est fondamental. C'est cependant un état qui ne s'atteint jamais parce que la vie est continuellement remplie de situations qui déséquilibrent la personne. La synergie de la personne lui sert donc à combattre l'entropie. Par conséquent la personne est, à court terme, plus puissante dans ses actes. À moyen terme, elle a une meilleure qualité de vie et à long terme une espérance de vie plus grande.

L'apprentissage est donc ce changement dans les fonctions de la personne tel que visé par les objectifs pédagogiques, amplifié par l'enseignement et constaté par l'évaluation. L'approche pédagogique préconisée recherche donc la congruence verticale (dans le temps) entre les objectifs, l'enseignement et l'évaluation et l'intégration horizontales parmi les objectifs, les méthodes d'enseignement et les techniques d'évaluation des fonctions (IAPCS).

## 2. Opérationnaliser les objectifs

Peu importe notre vision de la personne qui apprend, il convient d'opérationnaliser cette conception de manière à la rendre utilisable par les enseignants. C'est pourquoi, le reste de la théorie est l'opérationnalisation de la planification, de l'enseignement et de l'évaluation de l'apprentissage. On a pris l'habitude d'appeler "objectifs" les intentions de changements que l'on souhaite voir apporter chez une personne. Les objectifs expriment le résultat de ce changement. Leur formulation laisse cependant prévoir la nature des changements eux-mêmes.

**L'objectif s'adresse à l'imagination.** Le changement projeté, pour être attractif et significatif, doit être perçu comme un besoin, un manque, un vide ou une nécessité. Ils servent à projeter dans l'avenir une image du résultat. L'étudiant doit s'y voir comme une personne plus puissante et plus harmonieuse. **L'objectif doit être authentique et significatif pour motiver l'étudiant.** Le consentement de l'étudiant constitue l'arrimage de son choix avec l'attrait de l'objectif. **L'objectif implique un effort.** Si nous hésitons à nous lancer tête première à l'assaut d'un objectif, c'est souvent parce que l'on ne sait pas trop quelle somme de travail l'objectif oblige ou encore parce que l'ampleur du travail est trop considérable. C'est pourquoi, il faut expérimenter les objectifs pour connaître l'effort requis en terme de temps et de stress. L'objectif décrit l'effort et le temps nécessaires à sa réalisation. **Un objectif est une description intelligente des résultats attendus.** On rédige un objectif pour la personne qui doit apprendre. Elle doit donc comprendre les termes comme s'il s'agissait d'un contrat. C'est pourquoi l'on doit porter une attention particulière à la formulation de l'objectif. **Un objectif représente un apprentissage essentiel** (Turcotte, 1993). Ce que l'on apprend ne sert pas seulement à soi. En fait, ce que l'on apprend sert souvent de monnaie d'échange avec les autres. On peut négocier ce que l'on apprend. Le contenu sert donc d'expériences profitables à la personne comme à la société.

**Les objectifs favorisent aussi la gestion des activités d'apprentissage.** L'on doit savoir où l'on va pour planifier, enseigner et évaluer de manière professionnelle. Les objectifs sont donc cohérents avec les fins de l'éducation et les buts des programmes. En outre, les activités d'enseignement/apprentissage et d'évaluation doivent être conformes aux objectifs visés. La sélection des objectifs est une activité professionnelle importante. On ne peut pas tout enseigner aujourd'hui puisqu'il est impossible de tout apprendre. Il faut choisir. Il faut d'abord donner à l'étudiant ce dont il a le plus besoin pour acquérir et maintenir son équilibre personnel. Il faut ensuite choisir des expériences qui le font avancer vers le but fixé. Ces expériences doivent être graduelles. Il faut ensuite les pondérer en fonction de leur importance et du temps que l'on veut consacrer à leur atteinte.

**Plusieurs taxonomies et typologies nous servent dans la formulation des objectifs.** Nous présentons une méthodologie qui est conforme à l'idée que nous nous faisons de la personne et des approches que nous proposons quelques auteurs choisis. Nous utilisons les idées de Bandler et Grinder (1980), Eisner(1979), Torrance (1966), St-Pierre(1992), D'Hainaut (1986) pour formuler les changements de la fonction imaginative. Krathwohl (1964) nous aide à formuler des objectifs dans la fonction affective. Harrow (1975) nous facilite la tâche dans la formulation des objectifs de la fonction physique. D'Hainaut (1986) nous aide à formuler les objectifs de changements dans la fonction cognitive. Saha (1988) et Kohlberg (1969) nous inspirent dans la formulation des objectifs de la fonction sociale.

Un objectif est une activité humaine projetée qu'une personne particulière n'est pas encore capable d'accomplir. La compétence est cette activité accomplie au-delà d'un certain seuil de performance, constaté par une évaluation adéquate. L'apprentissage est ce changement dans la personne tel que le définit l'objectif. Ce dernier indique dans sa formulation le chemin à parcourir de sorte que l'objectif inclut la situation de départ de la personne qui veut changer. C'est donc en effectuant l'objectif lui-même qu'on l'apprend. L'objectif se confond alors à l'activité d'apprentissage elle-même.

D'une manière générale, le tableau suivant représente la congruence "objectif - enseignement - évaluation" à l'égard des diverses fonctions à changer chez l'étudiant (imaginative, affective, physique, cognitive et sociale). L'apprentissage de la fonction imaginative est alors formulé à l'aide d'objectifs de création, qu'il est amplifié par des modèles d'auto-enseignement et constaté par les techniques d'entretien non directif. On voit aussi que l'apprentissage de nature affective est formulé par les objectifs décrits en terme de choix personnels et autonomes et d'intensité des activités de l'étudiant et de sa participation, qu'il est amplifié par des enseignements axés sur la personne et constaté par des échelles d'appréciation. On voit, de plus, que l'apprentissage d'ordre physique est formulé par des objectifs décrivant les formes d'énergie produite par le corps en santé, qu'il est amplifié par des modèles axés sur l'entraînement et l'alimentation et constaté par des grilles d'observation et des listes de vérification. On voit également que l'apprentissage de la fonction cognitive est formulé par des objectifs d'acquisition de connaissances et d'habiletés cognitives, qu'il est amplifié par des modèles d'enseignement axés sur le traitement de l'information et constaté par des tests critériés. On voit, enfin, que l'apprentissage à caractère social est formulé en termes de relations avec les autres, qu'il est enseigné par des modèles sur les interactions sociales et constaté par des questionnaires d'opinion, des enquêtes, des entrevues et des sondages.

Les phases de l'apprentissage et les fonctions en changement

PHASES → ↓ FONCTIONS	OBJECTIFS	ENSEIGNEMENT	ÉVALUATION
<b>IMAGINATIVE</b>	Créer des images significatives	Modèles d'auto-programmation	Étude de cas Entretien non directif
<b>AFFECTIVE</b>	Choisir avec enthousiasme	Modèles axés sur la personne	Échelle d'appréciation
<b>PHYSIQUE</b>	Se régénérer pour fournir de l'énergie	Entraînement	Grille d'observation Liste de vérification
<b>COGNITIVE</b>	Traiter l'information pour produire des connaissances	Modèles axés sur le traitement de l'information	Test critérié
<b>SOCIALE</b>	Négocier avec les autres pour obtenir ce que l'on veut	Modèles axés sur l'interaction sociale	Entrevue dirigée Enquête Étude systématique

Un objectif est une description compréhensive du résultat du changement anticipé. Pour élaborer un objectif de manière opérationnelle et intégrée, il faut décrire ce qui lui est préalable, l'objet sur lequel l'étudiant devra travailler, les processus requis, un par fonction, le produit attendu, le seuil de performance, les méthodes d'enseignement choisies et le mode d'évaluation congruent avec ce qui précède.

**Les éléments préalables** à un apprentissage doivent être connus de manière à lier harmonieusement l'objectif avec ce qui précède ou avec ce qui a été appris antérieurement. Parfois, l'on présume qu'un apprentissage préalable a eu lieu et que ce ne soit pas le cas, ou bien que cet apprentissage soit oublié. Comme les préalables cognitifs et affectifs déterminent les apprentissages subséquents, il convient d'inclure ces éléments dans l'objectif.

**L'objet** est la situation dans laquelle on place l'étudiant avant de réaliser l'activité à apprendre. Le contexte, les circonstances et les données font partie de l'objet. **Le processus** est ce qui se passe dans la personne pendant la réalisation de l'objectif. Chaque fonction a son processus qui lui est propre mais chacune transforme l'énergie vitale en un produit spécifique nécessaire à l'harmonie de la personne. Chaque objectif comporte un processus dominant selon qu'il sollicite davantage une fonction plutôt qu'une autre. **Le produit** est la situation dans laquelle se trouve l'étudiant après avoir réalisé l'activité quand il manifeste ce qu'il a appris. L'objet et le produit peuvent être des particuliers, des classes, des opérateurs ou des structures qui proviennent d'un contenu particulier.

Le **seuil de performance** est la limite d'acceptabilité de la réponse, la pondération des éléments de réponse et le temps nécessaire à la réalisation de l'objectif. Il est aussi très instructif et utile pour l'élaboration du plan de cours d'indiquer la méthode d'enseignement utilisée pour atteindre l'objectif, car l'étudiant apprend le modèle d'enseignement en même temps que le contenu. Si l'objectif est la description « en compréhension » de ce que l'étudiant doit apprendre, le domaine en est la description « en extension ». Il faut donc indiquer dans l'objectif les règles pour engendrer et sélectionner les items qui vont constituer l'instrument de mesure. Enfin, il est aussi très intéressant de connaître la modalité d'évaluation utilisée pour constater son atteinte.

### 3. Organiser son enseignement pour favoriser le succès

Il s'agit de comprendre les principes du succès pédagogique et de les appliquer de manière à obtenir un apprentissage satisfaisant. Le professeur doit organiser son enseignement et travailler à la réussite de l'étudiant. Il doit donc se concentrer sur les facteurs qui déterminent l'apprentissage. Bloom (1979), dans sa "pédagogie du succès", affirme que trois facteurs déterminent l'apprentissage. Ce que sait l'étudiant avant d'entreprendre un objectif détermine 50% de son rendement. Comment il se sent, d'un point de vue affectif, expliquent un autre 25%. La qualité de l'enseignement explique encore 25%. Un enseignement de qualité peut compenser les deux premiers facteurs. Selon Bloom, un professeur fournit un enseignement de qualité quand il ou elle présente ce qu'il y a à apprendre de manière engageante, quand les étudiants participent activement à leur apprentissage, que cette participation intègre une évaluation formative (diagnostic et prescription) cohérente et que l'on reconnaisse leur travail de manière convenable. De ces quatre caractéristiques d'un apprentissage de qualité, la participation de l'étudiant est la plus importante.

Un apprentissage de qualité est efficace, durable, étendu, transférable, intégré dans les fonctions de l'étudiant et fait appel régulièrement aux catégories élevées des fonctions. Un apprentissage est efficace quand il s'effectue rapidement avec le minimum d'investissement. Cela se produit dans « l'éclair de génie », l'insight, l'intuition. On ne connaît pas de moyen infaillible de stimuler ce genre d'apprentissage, mais les conditions favorables sont les mieux rencontrés dans les modèles d'enseignement qui stimulent la créativité, par exemple la synectique, la stimulation artistique, les problèmes à résoudre et les questions de métacognition. Ces conditions créent par nature une inquiétude, une anxiété que le professeur se doit de contenir et non pas d'éliminer. Un exemple d'apprentissage inefficace est celui de la télévision. Nous sommes bombardés par une multitude d'informations que nous ne sommes pas en mesure de stocker dans la mémoire. Ce procédé est le plus coûteux et le moins rentable, à moins qu'il introduise les boucles de répétition, comme c'est le cas de la série « Sesame Street », mais le coût est grand. Un autre exemple est l'enseignement magistral en grand groupe (au-delà de 25 étudiants) quand le contenu à apprendre dépasse le nombre maximum que l'on peut retenir, et « digérer » dans le temps de l'enseignement.

Un apprentissage est **durable** quand il persiste dans le temps et non pas seulement pour un examen. On sait que les apprentissages efficaces sont les plus durables. On peut faire durer un acquis par la répétition et l'exercice même s'il a été long et pénible à jaillir. C'est ce que l'on appelle les événements de la mémoire à court terme. Par ailleurs, les souvenirs stockés en mémoire à long terme sont permanents et jaillissent de manière imprévue.

Un apprentissage est **transférable** quand il peut servir à d'autres apprentissages. Cette qualité est le mieux acquise quand les enseignants se concertent. La planification des apprentissages entre les divers programmes, les divers niveaux et les divers enseignants favorisent le transfert des acquisitions. Un contre exemple d'un

apprentissage transférable est celui de l'étudiant qui a appris à appliquer la règle du participe passé mais qui applique cette règle moins de 8 fois sur 10 correctement dans un rapport de philosophie ou de géographie.

Un apprentissage est **intégré** quand il fait appel à au moins à trois fonctions de la personne. On désigne ici les fonctions imaginative, affective, physique, cognitive et sociale. La personne se sent mieux respectée quand on sollicite toutes ses fonctions dans l'action à apprendre. Elle se sent concernée, interrogée plus intimement, plus « personnellement » dans l'action. Elle agit donc avec plus d'énergie. L'apprentissage est donc plus efficace, durable et transférable. Ne faire appel qu'à une ou l'autre des fonctions est une pratique tellement généralisée que l'on peut identifier la fonction visée simplement en nommant le cours. Par exemple, Mathématiques égale cognitif, Education morale égale sociale et affectif, Danse égale imaginaire, Éducation physique égale physique. Il faudrait donc un travail concerté important au niveau des programmes pour intégrer les fonctions de manière harmonieuse.

Un apprentissage est **étendu** (apprentissage horizontal) s'il est diversifié et qu'il s'applique à divers champs de la connaissance, divers contenus et diverses situations. Certains enseignements techniques, mathématiques et professionnels sont souvent présentés de manière obtuse. L'étudiant ne soupçonne même pas à quoi va lui servir ce qu'il apprend. Les liens que le professeur fait avec d'autres aspects de la vie amplifient l'apprentissage, surtout si cela touche l'étudiant personnellement. Enfin, un apprentissage de qualité fait appel aux **niveaux élevés** (apprentissage vertical) des fonctions de la personne. On désigne les deux derniers niveaux de chaque fonction, la résolution de problème et la méditation (I), l'organisation et la caractérisation (A), l'expression non verbale et verbale (P), l'exploration et la mobilisation (C), et le civisme et l'altruisme (S). Un enseignement qui ne s'en tiendrait qu'au rêve (I), à la réception (A), à la perception (P), à la reproduction de connaissances (C) et aux relations un à un (S) aurait une piètre formation. L'environnement que procure le contenu disciplinaire est important. Il ne s'oppose pas à l'étudiant, bien au contraire. Mais si la quantité de contenu augmente, les choses que l'étudiant peut faire avec chacune diminuent, c'est-à-dire que la qualité des expériences diminuent.

### 3.1 Les préalables cognitifs

Les **caractéristiques cognitives** de départ sont pour Bloom les principales connaissances, habiletés requises à l'apprentissage d'une tâche. Un étudiant aborde une nouvelle tâche d'apprentissage avec une certaine histoire comportementale et des apprentissages acquis antérieurement. Une partie de ce passé détermine la façon dont l'étudiant se comporte vis à vis de ce nouveau défi et, par conséquent, les résultats de l'apprentissage.

Les différences de niveau ou de vitesse dans l'apprentissage s'expliquent, pour une bonne part, par l'histoire de l'élève. En effet, si tous les étudiants disposent des préalables à l'accomplissement d'une tâche d'apprentissage particulière, la variation du niveau ou de la vitesse d'apprentissage sera moins importante que si les prérequis ne sont pas également maîtrisés par tous.

Les caractéristiques antérieures de l'étudiant influencent fortement la façon dont il apprendra plus tard et le temps ou l'aide dont il aura besoin pour surmonter les difficultés nouvelles. La variation de rendement des étudiants à la fin de l'année ou du semestre est fortement liée aux variations du rendement observé pour ces mêmes matières ou pour des matières apparentées avant le début de l'année ou du semestre. Il existe une corrélation positive entre les mesures de rendement à la fin de l'année et le rendement de l'étudiant plusieurs années plus tard dans la même matière. Les apprentissages proposés dans une année scolaire donnée supposent que d'autres ont été réalisés antérieurement. Il paraît difficile de concevoir une tâche d'apprentissage qui ne soit pas basée sur un apprentissage précédent.

Les caractéristiques cognitives de départ sont modifiables. L'apprentissage peut s'améliorer considérablement si les étudiants sont placés dans des conditions favorables. Les caractéristiques cognitives de départ peuvent expliquer jusqu'à la moitié (50%) de la variance des apprentissages réalisés.

### 3.2 Les préalables affectifs

Les **caractéristiques affectives** de départ constituent le second facteur de l'apprentissage. Les individus varient dans leur état affectif vis à vis des apprentissages à réaliser. Leur intérêt, leur attitude et la perception qu'ils ont d'eux-mêmes varient. Si les étudiants abordaient une tâche d'apprentissage avec enthousiasme et intérêt, l'apprentissage serait beaucoup plus facile et ils apprendraient plus vite et mieux que les étudiants peu motivés. Un

individu tend à aimer les activités dans lesquelles il croit pouvoir réussir ou a déjà réussi. La notion de succès implique qu'à la tâche antérieure et à la nouvelle tâche était lié un défi, voire une prise de risque. L'affect de l'étudiant peut être en partie déterminée par sa perception de la relation entre la tâche et un ensemble plus vaste de buts ou d'objectifs ultérieurs présents dans son esprit. S'il considère que la tâche va le rapprocher de ses buts, sa réaction affective sera probablement positive ( et vice versa).

Le sentiment de réussite influence les performances scolaires qui, couronnées de succès, renforcent ce sentiment. Le succès ou l'échec dans une matière scolaire détermine la manière dont l'élève considère la matière et son désir d'en apprendre davantage. Le succès ou l'échec peut déterminer la manière dont l'étudiant voit l'école et l'apprentissage scolaire et son désir de continuer ses études à l'université ou en dehors de l'école, à l'âge adulte. Avec le temps, l'étudiant tend à développer une vision de lui-même qui reflète sa perception de son adéquation dans les tâches d'apprentissage scolaire. Le degré de réussite scolaire influence la santé mentale de l'individu.

### 3.3 La qualité de l'enseignement

La **qualité de l'enseignement** est le troisième déterminant de l'apprentissage. Un enseignement de qualité est un environnement qui favorise l'apprentissage de tous les étudiants. Tout enseignement de qualité comprend les quatre phases suivantes. L'enseignement débute par une présentation claire de ce qu'il faut apprendre. L'étudiant connaît l'organisation, la forme, la fréquence, la signification et la variété des éléments. Suit cette présentation une participation active de l'étudiant. Il s'exerce, discute, travaille seul ou avec les autres, produit, répond et pratique. Le professeur observe l'activité de l'étudiant et lui donne un feed-back. Il s'agit d'un diagnostic et d'une prescription. On appelle aussi cette activité "évaluation formative". Enfin, l'enseignement efficace se termine par un renforcement adéquat, varié et adapté à chacun. L'apprentissage n'est efficace que s'il s'accompagne d'un tel renforcement. L'enseignement de qualité intègre harmonieusement ces composantes.

**Quels sont les effets escomptés de l'enseignement sur le rendement?** L'indice le plus net de la qualité de l'enseignement est le degré de participation de l'étudiant. L'utilisation de l'évaluation formative exerce un effet déterminant sur le rendement des étudiants. L'enseignement collectif ne convient qu'à une petite partie des étudiants. Il faut tendre à individualiser si l'on veut que la plupart des étudiants apprennent plus efficacement. On peut modifier la qualité de l'enseignement de manière à permettre à plusieurs étudiants d'atteindre un niveau de maîtrise qui modifiera par la suite leurs caractéristiques cognitives et affectives de départ. Les étudiants qui réalisent de bons apprentissages sont, en général, ceux qui reçoivent un enseignement de bonne qualité. La qualité de l'enseignement détermine 25% du rendement des étudiants.

Il y a donc des facteurs déterminants de l'apprentissage. L'enseignement n'est donc pas un art mais la gestion harmonieuse des caractéristiques cognitives et affectives de l'étudiant par une présentation claire de la tâche à apprendre, une participation active de l'étudiant, un feed-back judicieux et une reconnaissance de l'acquis adéquate. Ces facteurs se combinent entre eux pour permettre jusqu'à 90 % des étudiants de réussir leur apprentissage dans le temps prescrit.

### 4. Choisir un modèle d'enseignement

À la suite de cela, on choisit le modèle qui convient pour passer à l'action. Joyce et Weil (1986) nous proposent plusieurs modèles d'enseignement. Selon eux, les étudiants n'apprennent pas seulement les contenus des cours mais la manière dont on les dispense. C'est pourquoi, le professeur doit apparié chaque objectif avec un modèle qui lui est congruent. Chaque modèle nourrit plus particulièrement une catégorie d'objectifs.

### 5. Choisir son mode d'évaluation sommative

L'apprentissage n'existe pas encore tant que l'on ne l'a pas constaté par l'évaluation. C'est le principe même de la validité. **Face aux choix** qu'a une institution d'enseignement de considérer un apprentissage comme acquis ou non, l'évaluation est une argumentation en faveur de l'une ou l'autre alternative qu'il faut chercher à rendre toujours plus **crédible**. Il faut donc **structurer** logiquement l'évaluation selon les modèles les plus appropriés en respectant les **standards professionnels** propres à l'évaluation, tels l'objectivité, l'impartialité, la fidélité et la validité. L'évaluation doit aussi respecter les **normes sociales** acceptées dans l'institution où elle a lieu, comme le partage des responsabilités concerté, la justice et l'équité. Elle doit en outre respecter les **principes pédagogiques** de cohérence

entre les fins, les buts et les objectifs de l'éducation ainsi que de congruence entre les objectifs, l'enseignement et l'évaluation. On évalue pour informer l'étudiant de son progrès (évaluation formative) ou la société du succès ou de l'échec de l'étudiant (évaluation sommative). L'évaluation doit toujours se baser sur des critères que fournissent les objectifs et recueillir des informations avec des instruments congruents et critériés (Roid et Haladyna, 1982).

## 6. Identifier le domaine du test

*Un domaine peut contenir n'importe quel ensemble d'items clairement spécifiés* (Hively, 1974). L'ancêtre du domaine est «le tableau de spécification» par lequel on décompose l'objectif en activité taxonomique et en contenu. Cela sert à répartir le temps alloué pour un cours ou le nombre d'items dans un examen.

Il est intéressant cependant de générer un domaine d'une manière empirique et reliée de manière plus congruente avec l'objectif à enseigner et évaluer. Par exemple, un professeur de tir au pigeon d'argile pourrait simplement suivre et observer des chasseurs sur le terrain. Il pourrait noter tous les points de départ des pigeons et la direction de leur trajectoire qu'il a observés sur une grande période de temps. De même, un domaine pratique de lecture pourrait être constitué à partir de toutes les premières pages des journaux de la dernière année. L'approche logique est une alternative à la collecte empirique d'items typiques. On peut combiner, de manière systématique, par exemple les vitesses, les angles verticaux, les trajectoires, les points d'origine, etc., pour spécifier le domaine. De même, on peut regrouper tous les mots choisis sur la première page des journaux de la dernière année selon des catégories grammaticales, des règles phonétiques, et des fréquences d'apparition. Chacune de ces caractéristiques théoriquement importantes subdivise le domaine en parties plus petites de manière à ce qu'il soit plus aisé de choisir des questions d'examens plus représentatifs de l'objectif visé et de faire des diagnostics et des prescriptions. *Une liste de règles pour engendrer un ensemble d'items reliés d'items s'appelle une "forme d'items"*. Une façon commode d'élaborer des domaines est de demander quelles parties d'un item que l'on peut changer pour créer d'autres items qui mesurent le même objectif. On dresse la liste des éléments de rechange aux parties variables de l'item. Puis on peut rédiger les règles de substitution qui permettent à ceux qui élaborent des tests, ou même un ordinateur, d'engendrer cet ensemble d'items reliés. Si l'objectif est la définition en compréhension d'une activité à évaluer, le domaine est son extension que l'on peut représenter par une population d'items ou une banque d'items ou par les règles qui permettent de les engendrer. Par ailleurs, un test est un échantillon représentatif de cette population. Voici les étapes générales d'élaboration d'un domaine.

## 7. Élaborer les instruments de mesure

L'item à choix multiple présente plusieurs choix de réponses à l'étudiant. Celui-ci choisit parmi les propositions la ou les bonnes réponses. Les "mauvaises" propositions sont appelées leurres ou distracteurs. Ils permettent de couvrir beaucoup de contenu en peu de temps d'administration. L'effet du hasard dépend du nombre d'items. Les experts reconnaissent la très grande difficulté de mesurer les habiletés plus complexes avec des items à choix multiples justement parce que l'activité de choisir la bonne réponse interfère avec celle qui est mesurée.

Le test à réponse courte présente une série d'énoncés qui constituent l'ensemble-stimuli. Pour chacun des éléments de ce premier ensemble, l'élève doit associer un élément faisant partie d'un deuxième ensemble, l'ensemble-réponses. Tous les éléments doivent référer à une même structure logique ou factuelle. Ils servent à mesurer l'habileté de conceptualisation parce qu'ils ont la même forme que les présentations de concepts: exemples et contre-exemples. On peut leur construire des domaines à partir desquels on peut échantillonner les items et assurer ainsi la validité des tests. Il est rare que l'on puisse utiliser la forme écrite pour ce genre de tests parce que les exemples qui possèdent une caractéristique d'un concept ne sont que rarement pertinents quand ils sont représentés sur papier.

Comme D'Hainaut (1988) le définit, il y a **résolution de problème** lorsque au moins un des facteurs suivants est nouveau pour celui qui résout le problème: la classe de situation initiale, le processus de résolution et la classe de la situation finale (solution). D'Hainaut nous propose une des classes d'analyse détaillées de la résolution de problème. Il souligne que le principal facteur de cette activité est son caractère nouveau. St-Pierre (1993) nous suggère de définir les problèmes avec des experts du domaine. Il s'agit d'abord de classer le problème avec ces experts. Ce qui permet d'élaborer une échelle d'appréciation. La situation stimulante présente un problème à résoudre et le cadre de référence nécessaire à la solution.



On utilise l'échelle d'appréciation pour effectuer la critique d'art et juger de travaux artistiques. Pour Eisner (1979), la critique est essentiellement qualitative, comme l'oeuvre elle-même de l'artiste. La principale tâche est de retrouver les qualités essentielles d'une oeuvre et les reformuler de manière à les faire percevoir plus profondément par les autres. Les qualités que le critique décrit peuvent se trouver dans l'oeuvre elle-même. La valeur d'une critique réside dans l'effet qu'elle a sur l'audience. Eisner distingue entre l'activité du "connaisseur" et celle du "critique". La première est l'art de l'appréciation, tandis que la seconde est celle de distinction. L'activité du connaisseur consiste à reconnaître et apprécier les qualités d'événement particulier sans exiger un jugement public ou une description. C'est le prélude nécessaire à la critique. L'activité du connaisseur nécessite chez celui-ci une grande expérience pour lui permettre de distinguer ce qui est significatif au sujet d'une peinture, d'une pièce de théâtre, une interprétation musicale. Il n'est cependant pas suffisant de reconnaître et classer un événement à partir de son expérience. L'on doit aussi pouvoir percevoir les subtilités et comprendre comment elles contribuent à l'oeuvre. Il faut donc une formation spécifique.

On observe quand l'objectif suggère la réalisation d'une habileté dont les aspects et les qualités sont tangibles, par conséquent directement observables plutôt qu'abstraites. La fiche individuelle nécessite que l'observateur note par un crochet (✓) si une tâche a été complétée ou a été réalisée. Un groupe d'observations constitue une liste de vérifications. La fiche sous-entend, bien sûr, tout un développement systématique, y compris un essai du test, de manière à ce que le produit ou la performance à observer soit bien défini. La fiche doit correspondre de près à l'objectif de l'enseignement. L'observateur peut avoir à produire une description narrative de la performance ou du produit. Il doit donc produire au préalable un document très spécifique pour la correction. À partir d'une fiche bien définie ou d'une grille d'observation, la méthode d'observation peut varier. L'observateur peut-être présent et visible à l'étudiant, l'observateur peut être présent mais non visible et il peut être absent mais capable d'inférer indirectement que le produit a été réalisé ( par vidéo, par exemple). L'observation systématique peut être effectuée à l'aide d'une montre pour les mesures qui requièrent un temps de réaction, une vitesse, une force ou un comptage d'erreurs. La grille d'observation s'emploie quand on veut observer plusieurs types de comportements ou d'actions différents. On utilise le système de signes pour déterminer si le comportement est présent (+) ou absent (-) dans un cadre de temps donné. Les grilles d'observation sont en fait des listes de vérification.

## 8. Administrer l'instrument de mesure

La situation est donc celle de la classe où l'étudiant s'est engagé à changer quelque chose dans ses fonctions de manière à répondre à un déséquilibre. Le professeur lui a proposé des objectifs et des activités qui devaient favoriser le changement souhaité. Les instruments de mesure cohérents avec ces objectifs et activités ont permis d'obtenir de l'information sur ce changement. Cette information a pris la forme de données. Puisque le but recherché était de vérifier le changement, il est donc essentiel de choisir des stratégies qui mettent en évidence ce phénomène.

Habituellement, le professeur suit une stratégie fort simple. Il enseigne et administre un post-test après son cours. Ce plan rudimentaire comporte beaucoup de sources d'invalidité. Chaque fois cependant que cela est possible, il est judicieux d'effectuer un prétest avant l'apprentissage et un post-test après cet apprentissage. Cette stratégie est fort efficace pour déceler l'apport réel du cours à l'apprentissage. On peut ainsi identifier quatre groupes d'étudiants. Un premier groupe d'étudiants serait en situation de non maîtrise avant le cours et en situation de maîtrise après le cours. Un second groupe d'étudiants serait en situation de non maîtrise avant et après le cours. Un troisième groupe d'étudiants serait en situation de maîtrise avant et après le cours. Enfin, un quatrième groupe d'étudiants serait en situation de maîtrise avant le cours mais en situation de non maîtrise après.

L'identification de ces types de situations permet l'amélioration réelle des cours, des tests et des interventions. Par exemple, au cégep, une tendance veut que les cours de Philosophie et d'arts et lettres soient trop souvent du type C(1,1). Dans ce cas, les étudiants qui réussissent sont ceux qui maîtrisent les objectifs, déjà, avant de commencer le cours. Une autre tendance serait que certains cours de sciences systématisent trop souvent la situation B(0,0). Dans ce cas, il n'y a pas d'apprentissage suffisant dans les cours. La situation recherchée par tout professeur est que la plupart des étudiants (plus de 80%) ne maîtrisant pas les objectifs au début, ce qui est normal, atteignent plus de 80% des objectifs du cours: A(0,1). La situation où les objectifs avant le cours, mais qui ne les maîtrisent plus à la fin, ne devrait pas se rencontrer. La stratégie du prétest suivie du cours et du post-test comporte certaines lacunes importantes parce qu'elle ne permet pas de contrôler des éléments qui peuvent invalider les résultats. En utilisant cependant un groupe contrôle qui n'a pas suivi le cours, on obtient une bien meilleure crédibilité de ses données.

Un **item** est l'élément de base du test. Il comprend (1) une mise en situation présentée à l'étudiant sous forme de situation stimulante. On appelle aussi cette partie la question. La (2) seconde partie de l'item est la procédure selon laquelle l'étudiant va exprimer sa réponse. Ce peut être un crochet entre parenthèse, mais aussi un texte ou un graphique ou un dessin. Il faut donc prévoir une procédure pour recueillir cette réponse. La (3) troisième partie de l'item est la clé de correction. On doit prévoir la manière d'interpréter la réponse. On doit donc établir des critères de correction de manière à évaluer toutes les réponses avec objectivité et impartialité.

La **correction** des épreuves est fastidieuse. Il serait souhaitable de robotiser cette opération, par exemple en utilisant le lecteur optique, léger comme un stylo, avec des codes zébrés. L'emploi de l'ordinateur facilite ensuite la correction et le feed-back rapide à l'étudiant. Avec l'avènement des ordinateurs ultra légers et portatifs, il devient plus facile au professeur de corriger les réponses des étudiants sur-le-champ, c'est-à-dire en classe même.

Il faut utiliser un **système de notation** qui conserve la cohérence de l'interprétation avec l'objectif. Les notes de tests critériés s'interprètent mieux en proportion ou pourcentage d'objectifs atteints. Que l'on utilise des lettres ou des chiffres, cela n'a pas vraiment d'importance en autant que la cohérence est conservée et que les usagers interprètent tous de la même façon les résultats. Il faut se concentrer sur l'essentiel quand on évalue (Turcotte, 1993). Un groupe d'experts dans la matière peut aider à se concentrer sur les éléments essentiels d'un programme ou d'un cours.

Parce que l'évaluation débouche sur un jugement de succès ou d'échec d'un étudiant, il faut établir un **seuil de compétence**. L'établissement du seuil implique les valeurs des professeurs. Ce travail devrait être fait en collaboration avec des collègues experts pour minimiser certains biais personnels et pour harmoniser les seuils au sein d'un établissement. Ces juges devraient pouvoir établir le seuil en tenant compte de la difficulté du test, de son importance pour la suite des étudiants. 80% de réussite est souvent le seuil recherché pour exprimer la compétence alors que les institutions n'exigent que 60%. C'est pourquoi on fait la différence entre le seuil de réussite basé sur le concept de maîtrise minimale et le seuil de compétence.

## 9. Réviser l'instrument de mesure

Selon Hambleton et Roger (1991), Swaminathan (1975) et Crehan (1974) suggèrent d'utiliser les estimés de Carver (1970) pour évaluer la fidélité et la validité des tests de compétence. En effet, Carver propose de comparer la proportion de décisions consistantes, à partir de tests parallèles, sur la compétence de chaque individu d'un groupe étudié. Il est évident que le concept d'équivalence de Carver est d'importance primordiale dans une situation de mesure de la compétence et que, par conséquent, nous l'adopterons pour cette démarche. On considère donc comme consistant un test (c'est-à-dire, un échantillon d'items) qui classera chaque examiné de la même manière qu'un test parallèle (ou équivalent): compétent aux deux tests ou incompétent aux deux tests.

Le soin que l'on met pour soutenir la cohérence tout au long de l'élaboration du test entre la définition des objectifs et les items constitue la trame logique de la validité du test. On peut cependant s'interroger sur l'efficacité de ces précautions et chercher objectivement à corroborer cette argumentation par des analyses empiriques. Dans un modèle pédagogique, le rendement de l'étudiant doit être faible au début de son apprentissage et élevé à la fin. Les performances aux tests devraient confirmer cette hypothèse sur le rendement et on devrait s'attendre à ce que les moyennes des scores entre le début et la fin du cours soient très différentes. Des tests de différences devraient confirmer cette grande différence dans les scores aux tests. Si l'item est une partie analogue au test, on peut s'attendre au même type de comportement de la part de l'item.

Le premier indice que nous proposons est celui de Harris et Subkoviak (1986) que l'on utilise quand on n'a pas pu utiliser de prétest: le schème avec post-test. L'analyse porte donc sur le post-test. Il se calcule facilement au niveau de la classe. La logique qui sous-tend cet indice est que si un item est réussi par des étudiants compétents alors qu'il est échoué par des étudiants incompétents, on a de l'évidence que l'item est valide. Il mesure ce qu'il est censé mesurer. Le second indice est celui de Cox et Vargas (1966). On peut l'utiliser quand on a choisi le schème avec prétest et post-test. La différence dans les difficultés d'items avant et après un enseignement est un type de mesure de sensibilité pour les tests reliés à un domaine. Cet indice de sensibilité à l'enseignement mesure d'assez près et d'une manière directe les effets de l'enseignement. L'indice de sensibilité devient alors un indicateur du degré d'erreur de mesure de cet item. On bâtit d'abord le tableau de contingence suivant:

		ITEM (x) au post-test	
		Succès (1)	Échec (0)
ITEM (x) au pré-test	Échec(0)	A	B
	Succès(1)	C	D

- Où A est le nombre de sujets qui ont échoué l'item (x) au prétest et l'ont réussi au post-test;  
 B est le nombre de sujets qui ont échoué l'item (x) au prétest et au post-test;  
 C est le nombre de sujets qui ont réussi l'item (x) au prétest et au post-test;  
 D est le nombre de sujets qui ont réussi l'item (x) au prétest et l'ont échoué au post-test.

On obtient ainsi quatre patrons possibles pour l'analyse. A est le patron idéal. B est un item dont l'habileté n'a pas été apprise par les étudiants, ou un item trop difficile ou un item pour lequel l'enseignement n'a pas été efficace. L'interprétation se fait à la lumière de l'expérience. C est le patron où l'item est trop facile et l'enseignement inutile. D est le patron à éviter. L'indice de Cox et Vargas est la différence entre A et sur N (le nombre de sujets).

### Conclusion

La boucle est enfin bouclée. Nous avons terminé notre argumentation de congruence « objectif - enseignement-évaluation ». Le périple a débuté par une vision de l'être humain et de l'apprentissage. L'opérationnalisation a précisé et clarifié l'habileté à apprendre. Elle s'est traduit en stratégie et en action auprès des étudiants. L'évaluation a reflété fidèlement, validement, objectivement et impartialement le résultat de cette action. L'évaluation doit s'exprimer en fonction des intentions du changement à apporter chez l'étudiant. Le changement a-t-il eu lieu comme prévu? C'est le rôle de l'évaluation de fournir l'information nécessaire pour répondre à cette question. Toutefois, l'évaluation ne dit rien de plus que ce qu'elle est censée dire, mais elle doit tout le dire.

### Bibliographie

- BANDLER, R. ET GRINDER, J. (1982), *Les secrets de la communication*, Actualisation edh.
- BLOOM, B.S. (1979), *Caractéristiques individuelles et apprentissages scolaires*, Labor-Nathan, Coll. 2000.
- CHANGEUX J.-P. (1983), *L'homme neuronal*, Fayard.
- COX, R. C. & VARGAS J. (1966), *A comparison of item selection techniques for norm-referenced and criterion-referenced tests*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, April.
- D'HAINAUT, L. (1988), *Des fins aux objectifs de l'éducation*, Labor Nathan, Coll 2000.
- EISNER, E. (1979), *The Educational Imagination*, Harper.
- HAMBLETON, R. K. (1991), "Advances in Criterion-Referenced Measurement". in Hambleton & Zaal (Eds.), *Advances in Educational and Psychological Testing*, Kluwer Academic Publishers.
- HARRIS, D. J. & SUBKOVIK, M.J. (1986), "Item Analysis: A Short-Cut Statistic for Mastery Tests, *Educational and Psychological Measurement*, 46.
- HARROW, A. (1975), *La taxonomie des objectifs psychomoteurs*, McKay
- HIVELY, W. (1974), "Introduction to domain-referenced testing", *Educational Technology*. pp. 5 - 16.
- JOYCE, B. & WEIL, M. (1986), *Models of Teaching*, 3e éd., Prentice Hall, 1986.
- KOHLBERG, L. (1969), Stage and sequence: The Cognitive-developmental approach to socialization. *Handbook of socialization theory and research*. Chicago: Rand McNally.
- KRATHWOHL, D.R. & al. (1964), *Taxonomy of Educational Objectives. Handbook II: Affective Domain*, McKay.
- RACINE, S. (1994), *La personne synergique. Manuel de pédagogie*. (Sous presse), 151 pages.
- ROID G.H. & HALADYNA, T.M. (1982), *A technology for test-item writing*, Academic Press.
- SAHA, L. J. (1988), "Sociometric methods", In *Educational Research, Methodology and Measurement: An International Handbook*, J. P. Keeves (Ed.), Advance in Education, Permagon Press, 753 -755.
- ST-PIERRE, L. (1992), *Un modèle d'évaluation de la résolution de problèmes en mathématiques au primaire*, U de M, Thèse de Ph. D.
- TORRANCE, E. P., (1988), "The nature of creativity as manifest in its testing". In R.J. Sternberg (Eds.) *The Nature of creativity*, Cambridge University Press. pp. 43 - 75.
- TURCOTTE, C., (1993), *Un modèle d'évaluation de la compétence scolaire*, U de M, Thèse de Ph.D.