

LOUISE BALAUX

**LES ÉPREUVES SYNTHÈSE DE PROGRAMME AU COLLÉGIAL,
QUELQUES CONSIDÉRATIONS THÉORIQUES ET PRATIQUES**

Essai
présenté
à la Faculté des études supérieures
de l'Université Laval
pour l'obtention
du grade de maître ès arts (M.A.)

Département d'orientation, d'administration et d'évaluation en éducation
FACULTÉ DES SCIENCES DE L'ÉDUCATION
UNIVERSITÉ LAVAL

JUIN 1998

© Louise Balaux, 1998





BIBLIOTHÈQUE

Cégep de l'Abitibi-Témiscamingue
Université du Québec en Abitibi-Témiscamingue

Mise en garde

La bibliothèque du Cégep de l'Abitibi-Témiscamingue et de l'Université du Québec en Abitibi-Témiscamingue a obtenu l'autorisation de l'auteur de ce document afin de diffuser, dans un but non lucratif, une copie de son œuvre dans Depositum, site d'archives numériques, gratuit et accessible à tous.

L'auteur conserve néanmoins ses droits de propriété intellectuelle, dont son droit d'auteur, sur cette œuvre. Il est donc interdit de reproduire ou de publier en totalité ou en partie ce document sans l'autorisation de l'auteur.

RÉSUMÉ

L'épreuve synthèse de programme au collégial présente un défi nouveau, un élément majeur lié au renouveau de ce niveau d'enseignement amorcé au début des années '90. Le concept de l'épreuve demeure mal défini au moment de sa mise en application pour la sanction des études en 1999.

Une analyse de textes ministériels et d'écrits sur les nouvelles approches en évaluation des apprentissages et sur des expérimentations d'épreuves synthèses, surtout aux États-Unis, démontre la complexité de la mise en oeuvre de telles épreuves.

Cet essai tente de cerner, d'une part, par une revue de littérature en trois volets, le concept d'épreuve synthèse dans un contexte d'évaluation des compétences, évaluation en fin de programme d'études. D'autre part, il ouvre quelques pistes de réflexion utiles aux concepteurs des épreuves et aux personnes intéressées à la mesure et l'évaluation des apprentissages complexes au niveau des études supérieures.



Louise Balaux

Gérard Scallon

Jean-Jacques Bernier

Juin 1998

TABLE DES MATIÈRES

INTRODUCTION	1
Chapitre 1 LES ÉPREUVES SYNTHÈSE DE PROGRAMME ET LE RENOUVEAU DE L'ENSEIGNEMENT COLLÉGIAL	3
1.1 Cadre réglementaire du renouvellement de l'enseignement collégial	3
1.2 Historique	5
1.3 Les épreuves synthèse de programme dans les textes de référence du renouvellement	8
1.4 Des impacts et des répercussions	14
Chapitre 2 LES ÉPREUVES SYNTHÈSE DE PROGRAMME ET LES NOUVELLES APPROCHES D'ÉVALUATION	17
2.1 Les caractéristiques connues de l'ÉSP	18
2.2 Les compétences comme objet d'évaluation des apprentissages à l'ÉSP	22
2.3 Quelques nouvelles approches d'évaluation des apprentissages vs l'ÉSP	27
2.3.1 <i>L'assessment</i>	29
2.3.2 <i>L'assessment</i> authentique	31
2.3.3.1 Le portefeuille, un outil d' <i>assessment</i> authentique	34

2.3.3	Évaluation des compétences dans des tâches complexes	38
2.3.3.1	La planification d'épreuves dans un modèle centré sur les standards	40
2.3.3.2	L'ECOS, un exemple d'évaluation des performances complexes	43
2.4	Par où commencer?	45
Chapitre 3 QUELQUES VISIONS ET EXPÉRIENCES D'ÉPREUVES SYNTHÈSE		48
3.1	Recherche d'information	48
3.2	Classification de l'information	50
3.3	Analyse et découpage de l'information	50
3.4	Information complémentaire	54
3.4.1	Des avantages et quelques inconvénients de l'utilisation des épreuves	55
3.4.2	Des questions soulevées, des réflexions	57
CONCLUSION		59
BIBLIOGRAPHIE		63
Annexe A	ÉPREUVES SYNTHÈSE : CARACTÉRISTIQUES, EXPÉRIMENTATIONS, VISIONS	70

INTRODUCTION

Depuis la réforme de l'enseignement collégial québécois, au début des années '90, plusieurs éléments nouveaux ont fait réagir les acteurs dans ce milieu, tant les étudiants que les administrateurs et les enseignants. Entre autres choses, le passage d'un d'enseignement disciplinaire avec des pratiques d'évaluation normative traditionnelle, par cours, à une organisation des plans d'études cadres et des plans de cours basés sur les compétences, idéalement dans une approche programme, le tout chapeauté de l'exigence d'un examen de synthèse à interprétation critériée en fin de programme ne va pas sans peine et sans heurt.

L'enseignement basé sur les compétences, pour les programmes ou les cours où il est implanté, a déjà passablement chamboulé les habitudes et les pratiques. Toutefois, l'épreuve synthèse de programme a été encore plus discutée à l'égard de sa pertinence, au départ, et de sa faisabilité jusqu'à maintenant, par les instances syndicales et même administratives.

Avant de délivrer son diplôme d'études collégiales, le Ministère souhaite vérifier, en imposant cette épreuve en fin d'études, si l'étudiant a bien intégré les apprentissages prévus dans l'ensemble de ses cours et s'il est prêt à assumer un rôle de débutant sur le marché du travail ou dans ses études supérieures à l'université.

Dans les collèges du Québec, la réflexion est bien amorcée sur ces épreuves qui présentent un attrait certain pour plusieurs; toutefois, il n'existe pas de modèle de référence pour en guider la mise en oeuvre qui est sous la responsabilité de chacun des collèges. Le présent essai cherche à éclairer notre lanterne sur ce sujet neuf; il est structuré en trois volets. Dans un premier temps, nous brosserons l'historique qui a précédé l'instauration des ces épreuves (et la mise en place de la

réforme) et nous chercherons à identifier les visées poursuivies par la Ministère en imposant cette épreuve de fin d'études. Dans un deuxième chapitre, nous tenterons de cerner la nature de ce que pourraient être ces épreuves synthèse, d'identifier certaines de leurs caractéristiques et d'en apprécier la spécificité par rapport à d'autres modes d'évaluation des apprentissages. Pour ce faire, nous allons jeter un regard sur le concept des compétences même et sur les nouvelles approches d'évaluation comme l'*assessment* et l'évaluation des tâches complexes. Enfin, dans un troisième volet, nous analyserons quelques textes faisant état de visions ou d'expérimentations d'épreuves de synthèse.

L'analyse documentaire des chapitres deux et trois porte principalement sur les modes d'évaluation des apprentissages en fin d'études, fin d'année, d'étape, de programme, évaluation allant au-delà des cours mêmes et référant à la fusion des habiletés, à la mesure et à l'évaluation des compétences ou des apprentissages «intégrés» aux niveaux d'études supérieures.

Nous souhaitons que cette étude serve de banque de documentation commentée, notamment aux personnes intéressées à la conception et à l'élaboration des épreuves synthèse et au domaine de la mesure et de l'évaluation.

Chapitre 1

LES ÉPREUVES SYNTHÈSE DE PROGRAMME ET LE RENOUVEAU DE L'ENSEIGNEMENT COLLÉGIAL QUÉBÉCOIS

Pour quelques mois encore, l'étudiante ou l'étudiant du niveau collégial qui aura réussi tous les cours de son programme obtiendra son diplôme d'études collégiales (DEC). Cependant, à partir de janvier 1999, une exigence nouvelle s'ajoutera pour l'obtention du DEC, soit celle de réussir l'épreuve synthèse de programme. Actuellement, pendant la période d'expérimentation, les collèges sont tenus de faire subir cette épreuve à tout élève qui a été admis à un tel programme à compter de l'automne '94. Il s'agit là d'un élément nouveau, sans précédent et fondamental du renouveau de l'enseignement collégial amorcé officiellement depuis '93.

Pendant que, dans les collèges, on s'affaire à élaborer ces épreuves, regardons, dans un premier temps, l'éclairage que les écrits ministériels et gouvernementaux nous apportent sur ces examens de fins d'études. Quelques textes concernent directement les épreuves synthèse de programme communément nommées ÉSP; d'autres s'y rapportent indirectement mais présentent un certain intérêt pour justifier l'instauration de ces ÉSP ou pour en comprendre les finalités.

1.1 Cadre réglementaire du renouveau de l'enseignement collégial

Au printemps et à l'été '93, le Gouvernement du Québec adoptait deux nouvelles lois et un règlement qui ont amené un changement majeur de

l'enseignement collégial au Québec. Les repères législatifs et réglementaires qui décrivent l'esprit et les visées principales de ce renouveau sont au nombre de trois :

- *La Loi modifiant la Loi sur les collèges d'enseignement général et professionnel (CÉGEP);*
- *La Loi sur la Commission d'évaluation de l'enseignement collégial (CÉEC);*
- *Le Règlement sur le Régime des études collégiales (RRÉC);*

Avec l'abrogation de la loi sur le Conseil des collèges, l'ancienne Commission de l'évaluation a été abolie et la Commission d'évaluation de l'enseignement collégial a vu le jour. Elle reçoit son mandat directement de l'Assemblée nationale, mandat centré essentiellement sur l'évaluation des politiques institutionnelles et la mise en oeuvre des programmes d'études tel que décrit dans le texte sur la mission et les orientations de ce nouvel organisme gouvernemental (janvier 1994).

Le champ d'action de la Commission peut sembler plutôt restreint mais ses pouvoirs et son autonomie de fonctionnement sont relativement grands. Ainsi, elle évalue les politiques institutionnelles d'évaluation des apprentissages (PIÉA) et leur application, les politiques institutionnelles d'évaluation des programmes d'études (PIÉP) et leur application, la mise en oeuvre des programmes d'études établis par le ministre de l'Éducation et enfin les objectifs et standards des programmes établis par l'établissement. De plus sa liberté d'action lui permet d'élaborer des critères et des instruments d'évaluation, de faire une évaluation chaque fois qu'elle le juge opportun, de faire des recommandations aux établissements et au ministre, de rendre public son rapport d'une évaluation ...

Le Règlement sur le Régime des études collégiales, dicté par décret du gouvernement en application de la Loi sur les collèges, traite essentiellement de la gestion des admissions, des dispositions relatives à la structure et au contenu des programmes conduisant au DEC, des responsabilités académiques partagées entre le ministre et les établissements, de l'évaluation des apprentissages et de la sanction des études.

1.2 Historique

L'adoption des textes de loi précités a fait suite aux travaux de la Commission de l'éducation de l'Assemblée nationale tenue à l'automne 1992, sur l'avenir de l'enseignement collégial québécois et sur les objectifs qui lui sont propres. Cette commission parlementaire a tenu ses auditions après 25 ans d'existence des collèges au Québec pour plusieurs raisons. Les collèges occupent une position stratégique dans le système scolaire québécois entre le secondaire et l'université et une intervention sur le collégial amènera un réexamen des autres composantes du système éducatif. La société québécoise et ses besoins ont beaucoup changé et l'émergence d'un environnement mondial obligent à des mises à jour et des modifications des contenus et des pratiques de l'enseignement. Les cégeps assurant l'ensemble de la formation des techniciens sont au coeur des enjeux de la main-d'oeuvre qualifiée du Québec. De l'avis du ministère, faute de pratiques communes d'évaluation, l'opinion publique, via des sondages et des palmarès, s'interrogeait sur la qualité de la formation, les taux de diplomation, la durée des études et autres dans les collèges. Finalement, la Commission de l'éducation poursuivait une réflexion déjà commencée par la Fédération des cégeps et par le Conseil des collèges, organisme aboli depuis.

Nous avons tenté de suivre le fil conducteur de cette réflexion du Conseil des collèges dans ses écrits qui ont précédé l'adoption des textes de loi sur l'enseignement collégial afin de retracer les problèmes ou les besoins à l'origine de l'instauration des ÉSP.

Avant même la tenue du débat public sur l'enseignement collégial, la Commission de l'évaluation, dans la conclusion de son rapport de janvier '92 sur l'*Examen des politiques institutionnelles d'évaluation des apprentissages et de leur mise en oeuvre*, lançait l'idée (ou la réactualisait) d'instaurer des examens synthèse ou travaux synthèse à la fin de chaque cours (ou ensemble d'activités d'apprentissages) et, pour les programmes préuniversitaires, une forme d'examen synthèse pour mesurer l'atteinte des objectifs de chacun des programmes (ensemble intégré de cours). En plus de déplorer le manque de

standards clairement définis au plan national et le fait que l'acte d'évaluer reste trop souvent un acte privé entre le professeur et ses élèves, la Commission constate que les mêmes cours ne sont pas équivalents en ce qui a trait aux objectifs poursuivis, aux compétences développées et au contenu dispensé dans un même collège. Il va sans dire qu'il est impossible de comparer les résultats d'un collège à l'autre. Dans ce contexte, les examens synthèse favoriseraient indirectement la standardisation des objectifs et des programmes et assureraient certaine équivalence entre les contenus de cours.

Dans ce même texte, la Commission annonce l'élaboration et l'expérimentation d'un examen synthèse de programme en sciences humaines et d'un examen témoin disciplinaire en mathématiques. Le mandat d'élaboration de l'examen en mathématiques était déjà formulé à l'automne 1991, sous l'égide de la Commission de l'évaluation et un comité expérimentateur provincial constitué de professeurs de neuf collèges a travaillé au projet de mai '92 à novembre '93 (Foucault, 1993). Quant à l'examen synthèse de programme en sciences humaines, il n'a jamais été mis en oeuvre.

Plusieurs années auparavant, dans son document de 1978 intitulé *Les collèges du Québec*, le ministère de l'Éducation soulignait déjà «l'absence de mécanismes capables de mesurer avec quelque précision la qualité des établissements et de l'enseignement qui y est dispensé, non moins que celle des diplômes décernés» (p. 42). Cette réflexion sur l'évaluation au collégial amorcée dans les années '70 se poursuit dans la recherche prospective menée par le Conseil des collèges intitulée *L'Enseignement collégial : des priorités pour un renouveau de la formation* (1992). Dans ce document, on met en évidence l'insuffisance d'évaluation des programmes d'études en soulevant la question de l'intégration des apprentissages chez l'étudiant : «... on dispense des cours, on les évalue un à un comme des pièces détachées, sans évaluer les objectifs de programme, sans créer de liens entre les apprentissages» (p. 262). On poursuit en insistant sur l'évaluation des programmes « ... parce que les apprentissages acquis dans chacun des cours ne prennent tout leur sens qu'en fonction d'un programme d'études» (p. 267).

Dans un autre document du Conseil des collèges, celui de Allaire & Moisan (1993), il est aussi question d'évaluation synthèse des apprentissages mais dans le cadre des évaluations de programmes. Il est dit dans ce guide opérationnel, qu'une évaluation synthèse permet de répondre aux questions suivantes sur le degré de réalisation des objectifs poursuivis dans le programme :

«Quels objectifs ont été mesurés par les évaluations synthèse des apprentissages? Y a-t-il une congruence entre les objectifs poursuivis et les évaluations synthèse des apprentissages?»

Les auteurs poursuivent avec une réflexion sur l'évaluation synthèse. Celle-ci permet de vérifier si l'étudiant peut faire des liens entre les apprentissages réalisés et permet d'évaluer les connaissances de base et les apprentissages plus complexes développés, faisant ici référence aux habiletés transversales (synthèse, analyse, transfert etc.). Cette évaluation synthèse peut prendre diverses formes : examen synthèse (objectifs d'ordre cognitif), projet de fin d'étude (ensemble des habiletés), grille d'observation en stage d'intégration (habiletés et attitudes), rapport de stage (réflexion dans l'action et liens entre les apprentissages).

Ces observations portent sur des aspects pédagogiques entourant la préparation d'évaluations synthèse mais, en résumé, Allaire & Moisan (1993) croient essentiellement que des données descriptives, en réponse aux questions citées, permettent de vérifier la congruence entre la réalisation des objectifs généraux du programme et les évaluations synthèse, dans un contexte d'évaluation de programme. Leur texte publié dans la collection *Études et réflexions sur l'enseignement collégial* et commandité par le Conseil des collèges (Commission de l'évaluation) exprime les opinions des auteurs inspirés des résultats partiels d'une expérimentation d'évaluation de programme dans trois collèges du Québec.

En résumé, pour le Ministère et son organisme ministériel de l'époque, c'est-à-dire le Conseil des collèges, les préoccupations reliées à l'intégration des apprentissages, à l'équivalence entre les cours voire entre les collèges et à la

qualité des programmes offerts ne sont pas des interrogations nouvelles soulevées uniquement lors des audiences de la Commission de l'automne '92. Il semble bien cependant que ces problèmes sérieux, identifiés de longue date, ont mené, entre autres, à l'instauration des épreuves synthèse dans le renouveau de 1993. On peut constater aussi que la notion d'épreuve synthèse est apparue dans les écrits du ministère quelques mois sinon quelques années avant les audiences de la Commission d'éducation. C'est à l'issue des travaux de la Commission que l'exigence des épreuves synthèse de programme a été imposée.

1.3 Les épreuves synthèse de programme dans les textes de référence du renouveau

En consultant les documents d'accompagnement du renouveau publiés par le ministère de l'Enseignement supérieur et de la science ou par la Commission d'évaluation de l'enseignement collégial dans l'ordre chronologique de leur publication, on observe que le ministère donne relativement peu de directives concernant l'élaboration des ÉSP et que la Commission, dans la plupart de ses textes même ceux portant sur l'évaluation des programmes, fournit des éléments sur les caractéristiques et les visées des ÉSP.

- *Des collèges pour le XXI^e siècle, L'enseignement collégial québécois : orientations d'avenir et mesures de renouveau.*

Le texte de présentation du renouveau de l'enseignement collégial publié par le ministère en avril 1993 identifie l'évaluation des apprentissages et la sanction des études comme cibles stratégiques du renouveau dans une perspective de responsabilités accrues pour les établissements. Concernant les épreuves synthèse, il est écrit que les politiques institutionnelles d'évaluation des apprentissages, déjà élaborées et utilisées par les collèges depuis plusieurs années, devront être enrichies et prévoir «l'instauration d'une épreuve synthèse à la fin de chaque programme d'études ... (épreuve) sous la responsabilité des collèges et gérée selon les dispositions de sa politique

d'évaluation des apprentissages». Dans ce texte aussi, on mentionne qu'il est souhaitable que les collèges se regroupent pour mettre au point leur épreuves synthèse afin d'assurer un plus haut taux de fiabilité, de comparabilité, d'équité. On note ici une préoccupation envers l'équivalence interinstitutionnelle dont la Commission de l'évaluation faisait état plus tôt dans l'examen des PIÉA.

- *Règlement sur le Régime des études collégiales*

Le texte du Règlement (juillet 1993) contient les dispositions relatives à l'évaluation des apprentissages et à la sanction des études qui nous intéressent. Plus spécifiquement, des passages des articles 25, 27 et 32 concernent les ÉSP et l'attribution des diplômes, et sont libellés ainsi :

Article 25 Le collège adopte, après consultation de la Commission des études, une politique institutionnelle d'évaluation des apprentissages et s'assure de son application. La politique institutionnelle d'évaluation des apprentissages doit notamment prévoir (...) une procédure de sanction des études et l'imposition d'une épreuve synthèse propre à chaque programme conduisant au diplôme d'études collégiales dispensé par le collège afin de vérifier l'atteinte par les étudiants de l'ensemble des objectifs et des standards déterminés pour ce programme.

Article 27 L'apprentissage est évalué pour chaque cours et pour l'ensemble du programme auxquels l'étudiant est inscrit.

Article 32 Le ministre décerne le diplôme d'études collégiales à l'étudiant qui, selon la recommandation du collège qu'il fréquente a atteint l'ensemble des objectifs et des standards du programme auquel il est admis, a réussi l'épreuve synthèse propre à ce programme et, le cas échéant, a réussi les épreuves uniformes imposées par le ministre.

On retrouve donc trois conditions pour obtenir un diplôme d'études collégiales : réussir tous les cours du programme, réussir une épreuve synthèse de programme et réussir des épreuves uniformes imposées par le ministre. Ces conditions s'appliquent à tous les programmes de diplôme du secteur technique ou du secteur préuniversitaire mais pas aux autres programmes dispensés par les collèges, par exemple les attestations d'études accordées par les services aux adultes et aux entreprises. Selon ces trois articles du Règlement, le but principal de l'épreuve synthèse réside dans la sanction des études seulement. Toutefois divers acteurs du renouveau semblent lui adjoindre d'autres buts potentiels comme on le verra un peu plus loin.

L'article 25 mentionne que l'épreuve doit vérifier l'atteinte de l'ensemble des objectifs et standards du programme. Or, un autre élément fondamental du renouveau, la redéfinition des programmes par la méthode dite des compétences assortie de la détermination des objectifs et des standards, relève du ministère. Cette opération est enclenchée certes pour les composantes de la formation générale (commune, propre et complémentaire) mais cette vaste révision de l'ensemble des programmes tant du secteur technique que du secteur préuniversitaire ne sera pas complétée avant l'échéance prévue pour la mise en application de l'épreuve synthèse. À l'origine le calendrier de mise en oeuvre de cette nouvelle ordonnance s'étalait sur 3 ans. L'application optionnelle de l'épreuve était prévue pour le trimestre d'hiver 1996. À l'hiver 1997, elle devenait obligatoire sans sanction et en 1998, elle devait être obligatoire et requise pour l'obtention du diplôme d'études collégiales. Depuis, cet échéancier a été retardé d'un an.

Dans l'article 27, il n'est pas explicitement question d'épreuves synthèse mais il est sous-entendu que l'ÉSP évaluera l'apprentissage pour l'ensemble du programme. Quant à l'article 32, il indique assez clairement que l'épreuve synthèse doit être une épreuve individuelle.

Dans les articles 25 et 32 du Règlement, il est mentionné que l'atteinte des objectifs et des standards est une condition d'obtention du diplôme. Cependant dans l'article 25, ceux-ci semblent liés à l'épreuve synthèse lorsqu'il

est écrit: «une épreuve synthèse ... afin de vérifier l'atteinte ... de l'ensemble des objectifs et standards déterminés». Dans l'article 32 pourtant, ils semblent s'en distinguer, devenant alors deux conditions indépendantes : l'atteinte des objectifs et standards (doit-on sous-entendre ici dans chacun des cours?) et la réussite de l'épreuve synthèse. Y a-t-il une distinction à faire? Aucun texte consulté jusqu'à ce jour ne clarifie cette question.

- *L'Évaluation des politiques institutionnelles d'évaluation des apprentissages - Cadre de référence*

Tel que mentionné plus haut, la Commission de l'évaluation de l'enseignement collégial est un organisme externe au ministère. L'étendue de ses pouvoirs et de son autorité, bien que définie dans son texte sur sa mission et ses orientations est encore mal connue.

En évaluant les PIÉA, la Commission a droit de regard sur la définition et les modalités d'application de l'épreuve synthèse. Dans son cadre de référence concernant l'évaluation des PIÉA, la Commission émet une vision un tant soit peu plus explicite que le RRÉC sur l'épreuve synthèse (janvier 1994, p. 13). Elle «considère cette épreuve comme un moment d'évaluation *dissocié* de l'évaluation des compétences acquises cours par cours; cette épreuve vise essentiellement à attester de l'intégration des apprentissages réalisés dans *l'ensemble* du programme» (le caractère italique est de la Commission).

Cette modalité de passation de l'ÉSP nous porte à croire d'une part que cette épreuve ne peut se situer à l'intérieur d'un cours même terminal et d'autre part qu'elle doit considérer, dans sa planification du moins, les apprentissages faits à l'intérieur de tous les cours du programme autant ceux de la formation générale que ceux de la formation spécifique. Dans le même texte, on ajoute par ailleurs que «l'épreuve peut prendre des formes variées et être située dans une activité d'apprentissage en fin de programme». Y a-t-il là contradiction? Généralement les activités d'apprentissage prennent place à l'intérieur d'un cours dans lequel des apprentissages spécifiques à ce cours sont évalués et

l'ÉSP doit mesurer l'ensemble des apprentissages du programme (ensemble de cours différents) y inclus ceux des cours terminaux.

La Commission d'évaluation élabore un peu plus sa définition de l'épreuve synthèse en précisant que celle-ci «prend en compte les objectifs et les standards déterminés par le ministre et le profil de sortie conséquent déterminé par l'établissement pour les futurs diplômés et les actions envisagées pour rechercher l'équivalence interinstitutionnelle». Qu'advient-il de l'équivalence interinstitutionnelle si, pour un même programme dans deux collèges différents et donc ayant les mêmes objectifs et standards éventuellement déterminés par le ministre, on a des profils sensiblement différents parce que déterminés par l'établissement? Comme l'écrit D'Hainault (1985) qui semble être un pionnier à définir la notion de profil de sortie, «le contenu de ce profil est déterminé par les options et les intentions de la politique éducative» de l'institution.

Toujours dans son cadre de référence pour l'évaluation des PIÉA (p. 19), la Commission expose les critères d'évaluation de l'application de la politique même. Entre autres, le critère de l'efficacité de l'application permettra de juger des dispositions prévues et actualisées de la politique qui garantissent la qualité de l'évaluation par «des modes et des instruments d'évaluation pertinents, cohérents, efficaces et transparents». On réfère alors bien entendu à l'évaluation des apprentissages, aux épreuves synthèse, aux examens, aux travaux etc.

Un autre critère d'évaluation des PIÉA, celui de l'équivalence de l'évaluation des apprentissages s'applique d'abord aux «mécanismes prévus pour assurer l'équivalence intra-institutionnelle de l'évaluation des apprentissages». On se basera par exemple sur les seuils de réussite, l'application de critères de correction, les exigences, la pondération des notes pour différents groupes d'étudiants ayant suivi le même cours, pour les cours d'un même programme ou pour l'ensemble des cours du collège. Toutefois, pour ce qui est de l'équivalence interinstitutionnelle, la Commission est muette quant aux modalités d'application du critère d'équivalence mais elle mentionne que son

application se limitera aux épreuves synthèse touchant des programmes dispensés par plusieurs collèges.

- *Guide général pour les évaluations de programmes d'études réalisées par la Commission d'évaluation de l'enseignement collégial*

Dans une optique d'évaluation des programmes, cette fois, la Commission, dans son Guide général (mai 1994 p. 19), replace à nouveau les épreuves synthèse au centre d'intérêt sous le critère numéro 5 dont le thème est l'efficacité du programme. Plus précisément, le sous-critère 5.5 qui concerne les standards dans l'acquisition des compétences établies pour le programme comporte deux exemples de questions d'appréciation éventuellement utilisées par la Commission en évaluation de programmes :

«Dans quelle mesure l'épreuve synthèse administrée par l'établissement permet-elle d'attester que les étudiantes et les étudiants qui terminent le programme satisfont aux standards établis pour mesurer l'atteinte des compétences visées?»

«L'établissement a-t-il développé des moyens en vue de s'assurer que le calibre de son épreuve de synthèse est équivalent à celui d'autres collèges?»

À la lumière de ces données, on constate que la Commission a manifestement l'intention de jeter un regard sur les épreuves synthèse comme mode ou instrument d'évaluation des apprentissages et comme élément permettant d'évaluer l'efficacité, la comparabilité et la transférabilité d'un programme d'étude. L'épreuve synthèse de programme devient alors un outil de choix dans un processus d'évaluation de programme en plus d'évaluer les apprentissages et de servir à la sanction des études.

À ce jour, dans les évaluations de programmes qu'elle a commandées depuis le début de son mandat (évaluations des programmes de technique de garde et technique informatique en '94-'95 et de sciences humaines en '95-'96), la Commission n'a cependant pas questionné ces programmes sur leurs

épreuves synthèse puisque celles-ci n'étaient pas encore au point pour l'application facultative en mai '96.

1.4 Des impacts et des répercussions

L'évaluation qui sanctionne les diplômes constitue un fait social à grande portée. Les épreuves synthèse de fin de programme feront partie de cette pratique et dès lors devront être pourvues de qualités rigoureuses (validité, fidélité, équité, équivalence).

Vraisemblablement, pour réussir ces épreuves, les élèves devront démontrer qu'ils ont intégré les apprentissages auxquels chacun des cours de leur programme les aura conviés. Cette nécessaire intégration des apprentissages aura des impacts majeurs sur la préparation et la prestation des cours par les enseignants.

De plus, les épreuves synthèse auront un impact sur les programmes puisqu'elles doivent mesurer l'atteinte des objectifs de ceux-ci et non pas seulement les objectifs disciplinaires de chacun des cours en faisant partie. Un effort particulier pour améliorer la cohérence et la qualité des programmes devra donc être déployé et l'ensemble des enseignants n'aura d'autre choix que d'adhérer au processus d'amélioration de la facture des programmes.

La Commission d'évaluation de l'enseignement collégial note, dans un rapport synthèse (janvier 1996, p. 9), que «des efforts importants ont été faits et des progrès ont été réalisés en vue de définir le cadre général de cette épreuve; des lieux ou des modes de concertation ont été mis en place afin de réaliser l'approche programme dont l'épreuve synthèse est largement tributaire».

Si ces épreuves synthèse impliquent la plupart ou toutes les matières du programme d'études elles forceront les professeurs à se concerter dans l'élaboration de leurs plans de cours et à réviser sinon améliorer leurs pratiques d'évaluation. La recherche de l'équivalence entre les épreuves

synthèse des différents collèges, telle que souhaitée par la nouvelle Commission d'évaluation de l'enseignement collégial, occasionnera indirectement l'équivalence entre les cours dans un même collège et entre les institutions.

Cependant l'obligation de préparer des ÉSP pour tous les programmes du collégial dans un délai relativement court génère une bonne dose d'anxiété et de frustration chez les responsables. Le manque d'information, de support et d'encadrement venant du Ministère dans cette démarche explique, partiellement du moins, cette situation. Les enseignants et les professionnels de l'enseignement collégial s'interrogent, à juste titre, sur la pertinence de ces épreuves, sur l'implication de toutes les disciplines enseignées dans le programme, sur la rédaction des objectifs et des standards liés aux compétences, sur la mise en oeuvre des instruments de mesure, sur l'administration des épreuves, sur les règles de sanction des études, sur l'intégration des apprentissages et la préparation des élèves à cet examen, sur la gestion des cas d'échec et des reprises etc.

Le fait que peu de programmes soient redéfinis en termes de compétences à ce jour pose un dilemme aux concepteurs des épreuves synthèse. Devront-ils décrire eux-mêmes leurs programmes en termes de compétences avant de préparer leurs épreuves synthèse? C'est une entreprise d'opérationnalisation de taille. L'absence de concertation entre les collèges dans cette démarche compromet la correspondance entre les diplômes et la comparabilité souhaitée entre les collèges telle que mentionnée par la Commission. Les collèges devront-ils plutôt élaborer des épreuves synthèse basées sur les objectifs très généraux actuels de plusieurs de leurs programmes? Si tel était le cas, quelle serait alors la valeur du diplôme accordé par le ministre ou, selon les règles et le langage du nouveau Régime pédagogique, le ministre décernera-t-il un diplôme pour la réussite des cours ou d'une épreuve synthèse d'un programme dont il ne connaît pas les objectifs et standards? De plus, la tâche d'élaboration de ces épreuves incombe aux professeurs du réseau dont les croyances et les pratiques sinon la compétence en matière de mesure et évaluation présente des lacunes comme l'ont décrit Howe et Ménard (1993) dans les résultats de leur recherche récente.

Dans son cadre de référence pour l'évaluation des PIÉA (janvier 1994, p. 12), la Commission d'évaluation demandait aux collèges de «déterminer au plus tôt les principales modalités d'application (*des épreuves*) afin que les étudiants admis en août '94 et qui y seront soumis puissent en connaître les principaux paramètres». Entreprise risquée s'il en est. En effet, comment peut-on identifier a priori les paramètres d'une épreuve synthèse de fin de programme quand on n'en connaît à peu près pas la nature? Qui plus est, cette épreuve devrait être équivalente aux épreuves synthèse du même programme dans les autres collèges.

Quatre ans après cette incitation à procéder avec diligence, des responsables de la conception des épreuves sont toujours en réflexion sur la nature des instruments en question, sur l'administration des épreuves et sur la sanction des études. D'autres tentent d'ébaucher des épreuves de qualité dont les résultats assureront, au mieux, une sanction juste et équitable des études pour chaque élève, selon le calendrier imposé. Après plusieurs mois de discussions et de représentations par les enseignants et les administrateurs, le Ministère a finalement accepté de reporter d'un an, en janvier '99, l'obligation d'imposer et de réussir l'ÉSP.

Finalement, les épreuves synthèse et les résultats des étudiants à ces examens permettront la sanction des études certes, mais sont aussi entrevus comme des outils privilégiés d'évaluation de l'efficacité d'un programme, évaluation qui elle-même, à moyen terme, «conduit à l'évaluation de l'établissement qui le dispense» tel que mentionné dans le texte du Conseil des collèges sur le renouveau de la formation (1992, p. 267).

Chapitre 2

LES ÉPREUVES SYNTHÈSE DE PROGRAMME ET LES NOUVELLES APPROCHES D'ÉVALUATION

La perspective d'évaluer, à l'aide d'une épreuve synthèse, l'ensemble des connaissances et habiletés acquises par un élève en fin de programme semble séduisante. On aura enfin une meilleure description du «produit fini» ou selon le discours actuel on pourra attester des compétences de l'élève sortant au collégial. Mais que signifie épreuve synthèse? Quelle peut être la nature de cette épreuve? Quelles réalités ou habiletés cerne-t-elle? Comment permet-elle de mesurer les apprentissages? Dans quelle situation place-t-elle les élèves? Voilà beaucoup de questions auxquelles la littérature ne fournira malheureusement pas toutes les réponses.

Quoique récemment réapparu au Québec, le concept de l'épreuve synthèse est loin d'être inédit. Déjà au 13^e siècle en Europe, l'étudiant universitaire était évalué oralement sur l'ensemble de ses connaissances théoriques et sur l'exégèse de ses expériences passées pour accéder à une profession. Ce format purement dialectique de l'évaluation demeura inchangé jusqu'au 19^e siècle lorsqu'à Harvard on introduisit un examen écrit et un système de gradation en quatre points, du succès à l'échec, dans une tentative pour régulariser et standardiser les exigences de graduation (Manus & al., 1992).

Aux États-Unis, l'utilisation de l'épreuve synthèse a connu des hauts et des bas depuis l'implantation du système de cours facultatifs qui a rendu les curricula fixes désuets, à la fin du 19^e siècle. Depuis ce temps les examens de synthèse ont été plus utilisés dans les Private Liberal Arts Colleges que dans

les autres institutions. Depuis une vingtaine d'années surtout, outre la fonction d'attestation des connaissances, on utilise parfois les résultats à des fins d'évaluation de la qualité d'un programme (Banta & Schneider, 1986).

2.1 Les caractéristiques connues de l'ÉSP

Reprenons maintenant quelques caractéristiques de l'ÉSP telles que véhiculées à date dans le milieu collégial en y ajoutant quelques notions auxquelles nous ramènent ces termes dans le but de mieux cerner les fonctions de l'épreuve et ses limites. Telle qu'imposée pour les programmes collégiaux, l'ÉSP est une épreuve sommative, terminale, individuelle, contribuant à certifier ou attester les compétences de l'étudiant.

D'abord l'aspect sommatif : selon la conception de Bloom (1971), l'évaluation sommative peut servir à d'autres fonctions que la certification telles que, comparer le rendement de groupes différents, prédire la réussite dans une formation, fournir un feedback aux élèves, faire le point avant d'entreprendre une nouvelle formation ou entrer sur le marché du travail, etc.

Cependant certaines de ces fonctions attribuées à l'évaluation sommative par Bloom apparaissent peu conciliables parce que trop différentes : par exemple, on ne peut prédire et certifier à la fois; ces deux fonctions sont incompatibles (Cardinet, 1988). Dans une visée d'évaluation passée et certificative, une seule question est posée sur un comportement terminal : l'étudiant est-il capable ou non de faire telle action? (Est-il compétent? Atteint-il les standards?) Aucune gradation dans la réponse n'est attendue. Dans une visée d'évaluation prédictive, de multiples questions peuvent être posées sur des aptitudes en corrélation avec un rôle futur. À l'aide d'examens élaborés en conséquence et mesurant ces aptitudes, on peut prévoir une gradation des résultats qui permette de distinguer les étudiants entre eux.

De même, on ne peut sanctionner et fournir un feedback utile du même coup. En sanctionnant, on émet le résultat d'une décision. L'étudiant obtient son diplôme ou ne l'obtient pas. Pour ce qui est du feedback en évaluation, il n'est utile que s'il informe sur l'assimilation de la matière à apprendre, sur des faiblesses, des besoins particuliers, des ajustements à faire, des pistes de solutions. L'évaluation sommative fournit, à la limite, des informations aux administrateurs ou à l'enseignant, pas à l'étudiant. Dans chacun de ces cas, les instruments d'évaluation devraient être différents parce qu'ils auraient des buts très distincts.

Le Ministère circonscrit l'ÉSP à la fonction certificative, certification individuelle et appelée externe par Tourneur (1985), certification dont la demande vient du système social externe à l'institution. La Commission d'évaluation y ajoute potentiellement la fonction d'évaluation de programme.

Au delà de ces considérations, il convient tout de même de redire que l'on envisage ici une épreuve ou une évaluation qui comporte quatre caractéristiques fondamentales soit une épreuve sommative, terminale, individuelle et certificative. Au moment de faire des choix et dans le contexte actuel d'élaboration des ÉSP, ces caractéristiques entraînent des implications ou à tout le moins quelques éléments de réflexion dont voici quelques exemples :

Épreuve sommative

- Ce terme nous réfère à l'**intention** de l'épreuve. Selon Bloom à qui cette nomenclature est empruntée, cette forme d'évaluation, dans le sens qui nous intéresse, sert à faire le point ou à totaliser les acquisitions passées.
- Le total des acquis représente une performance supérieure à la somme des parties. C'est là que le concept large des compétences prend une partie de son sens. Il ne s'agit pas d'évaluer des apprentissages superposés mais plutôt intégrés dans un savoir faire particulier.

- Elle diffère de toute évaluation formative; elle ne participe pas, comme le fait cette dernière, à la construction des savoirs, savoir-faire et savoir-être.

Épreuve terminale

- Cette terminologie nous reporte au **contenu** et au **moment** de l'épreuve.
- Le contenu représente un comportement terminal. Selon Cardinet (1988) le contenu de tests sommatifs et terminaux est «constitué de tâches globales» ... et dans ce cas, «l'accent doit être mis sur la finalité et donc sur le résultat de l'opération et non plus sur le processus».
- Dans le cas de l'ÉSP, l'épreuve a lieu au terme de tous les cours du programme.
- Elle porte sur le passé, l'ensemble des apprentissages.
- Elle ne tient pas compte de l'évolution des apprentissages dans le temps.
- Elle réfère au profil de sortie du programme.

Épreuve individuelle

- L'épreuve est individuelle parce que sa **visée** est la sanction des études. L'élève est donc l'auteur, le seul responsable de sa performance.
- Afin de comparer les performances des étudiants aux standards reliés aux énoncés de compétences, on devra développer des ÉSP dont l'interprétation des résultats sera inévitablement critériée.
- Cette visée exclut certains types d'épreuves, comme les productions de groupe ou impliquant l'aide de pairs.
- La décision qui découlera des résultats concerne l'individu. Dans un système de décision «succès-échec» qui prévaut habituellement avec l'interprétation critériée, il est concevable qu'une tâche ou un ensemble de tâches dans une épreuve puissent être accomplies avec excellence ou juste adéquatement par rapport au seuil de performance imposé. Ce système permet une sanction juste des performances individuelles mais donne peu d'information sur ces performances et pas d'information significative par rapport au groupe ou au programme.

- On ne cherche pas à distinguer les performances supérieures. Dans ce cas il ne devrait pas y avoir d'exigence compétitives ni de limite de temps pour résoudre un problème sauf si on a des raisons objectives d'en imposer.
- La détermination des seuils de performance, après l'élaboration des critères et des standards, est une opération délicate:
 - l'étudiant doit-il réussir à tous les critères listés?
 - certains critères sont-ils plus importants que d'autres?
 - doit-on exiger la réussite de critères dits essentiels et d'un certain pourcentage de critères dits pertinents?
 - les seuils exigés décrivent-ils la compétence minimale attendue par les enseignants ou la compétence souhaitée par des employeurs ou autres?
 - ces seuils sont-ils connus des étudiants?
- Parce que ce système d'évaluation critériée ne fournit aucune mesure relative entre les étudiants, les groupes ou les institutions, les résultats ne pourront être utilisés pour des fins d'évaluation de programme ou de comparaison entre les collèges. Seule l'épreuve et ses caractéristiques propres pourront être regardées en terme de validité ou autres qualités techniques, en terme de contexte de passation, de cohérence entre les critères et les standards du programme, de concordance entre les juges s'il y a lieu, etc.

Épreuve certificative

- Cette caractéristique fondamentale décrit la **fonction** de l'ÉSP.
- L'épreuve atteste des compétences à la sortie de tel programme dans un cégep donné; cela implique une préoccupation pour la validité de l'épreuve. Une question embarrassante peut se poser ici : Comment interpréter qu'un étudiant réussisse tous ses cours au premier essai et échoue éventuellement l'épreuve synthèse de programme?
- Elle représente une exigence partielle pour l'attribution de diplôme.
- Elle certifie uniquement que l'étudiant a atteint les objectifs ou des standards de performance précis.

- Finalement, la fonction certificative exclut l'utilisation de l'autoévaluation.

2.2 Les compétences comme objet d'évaluation de l'ÉSP

L'ÉSP doit mesurer l'atteinte des objectifs et des standards par l'élève à la fin de son programme et tel que mentionné précédemment, ces éléments ne sont pas clairement définis pour la plupart des programmes. Cependant le renouveau de l'enseignement collégial est bel et bien centré sur le développement des compétences et donc l'intégration des apprentissages.

Vers la fin des années '70, Spady (1980), principal artisan de la *Competency Based Education* (CBE) et par la suite de l'*Outcome Based Education* (OBE) offrait une conception élargie de la compétence et mentionnait quelques implications importantes de cette vision. Il décrivait alors essentiellement les compétences comme des indicateurs de la réussite ou de la performance dans les rôles et les activités de vie courante. (Dans la plupart des ouvrages américains consultés, on ne distingue pas explicitement les termes *competence* et *competency*). Cette courte définition implique cependant des éléments intéressants sinon importants qui résument ici, en cinq points, la pensée de Spady qui peut encore inspirer les responsables de l'élaboration des épreuves ou des programmes en terme de compétences et qui laissent entrevoir l'ampleur et la complexité de la mise en oeuvre des ÉSP :

1. Le focus et le contexte des compétences sont la vie courante et les rôles variés que nous occupons, lesquels requièrent un large éventail de capacités, capacités de performer en tant que consommateur, parent, partenaire, citoyen etc. Cela signifie que les curricula centrés sur le développement des compétences doivent prendre en compte, au départ, les demandes et les exigences de ces rôles majeurs et non pas la logique et la substance des sujets disciplinaires. Par exemple, dans cette optique, aucun de ces rôles fondamentaux ne s'appelle les mathématiques, les arts créatifs ou l'ingénierie.

2. Le succès dans les activités de vie courante exige de transiger avec les conditions sociales et les réalités très changeantes. Un des attributs essentiels d'une personne généralement compétente sera l'adaptabilité.
3. Les compétences sont formées à travers une intégration complexe de plusieurs capacités distinctes. Certaines de ces capacités sont évidemment très apparentes et mesurables; d'autres sont extrêmement subtiles ou même invisibles pour plusieurs. La compétence implique rarement la simple application mécanique de capacités cognitives ou manuelles.
4. Pendant que les connaissances, les habiletés et les concepts sont des composantes importantes du succès dans les rôles fondamentaux, ils ne l'assurent pas. La performance dans ces rôles est au moins également facilitée par les attitudes, les valeurs, les sentiments, les attentes, la motivation, l'indépendance, la coopération, l'endurance et l'intuition que les personnes possèdent. Les capacités affectives ne peuvent demeurer implicites dans un programme centré sur les compétences dans la vie courante.
5. Les compétences, en d'autres termes, reflètent à la fois ce que la personne est et ce qu'elle peut faire. Les programmes devront alors développer des outils d'*assessment* qui mettent le focus sur les aspects plus qualitatifs de la performance aussi bien que sur les démonstrations d'habiletés cognitives et manuelles observées par des mécanismes de mesure conventionnels. (Nous reparlerons plus en détail du terme *assessment* un peu plus loin, dans la troisième partie de ce chapitre).

Plus récemment, comme Perrenoud (1995) le propose, les compétences réfèrent «à des savoirs de haut niveau, qui exigent l'intégration de multiples ressources cognitives dans le traitement de situations complexes». Au sujet des connaissances, il ajoute «qu'elles sont des ressources cognitives souvent essentielles dans la construction d'une

compétence». Connaissances pratiques, connaissances scientifiques, connaissances d'expérience, connaissances savantes, les compétences ne sauraient se développer sans elles.

Pour sa part, Goulet (1995) soulève l'ambiguïté reliée au fait que deux définitions très distinctes des compétences risquent de créer une impasse dans le développement de cette approche par compétences au Québec. Une première vision étroite «présente les compétences comme des performances désirables, des tâches scolaires, des comportements reliées à des fonctions de travail spécifiques ... des objectifs de cours ...». Dans une deuxième conception plus large, les compétences sont considérées comme «des capacités de niveau supérieur, de résolution de problèmes, d'intégration de savoirs, de savoir-faire, de savoir-être» impliquant le transfert des apprentissages dans des situations non limitées au travail mais «touchant la vie personnelle et sociale».

Dans une autre tentative de clarification de ce concept clef, Désilets & Brassard (1994) présentent, dans un sens générique, une définition de la compétence comme «un état ou une qualité globale de la personne qui résulte de l'intégration appropriée des savoirs, savoir-faire et savoir-être pertinents à un domaine professionnel» en insistant sur le fait que la compétence «est un état de la personne et non pas une action ou un comportement». Bien que cette définition implique, dans les faits, un ensemble restreint de compétences spécifiques à un domaine, les auteurs appuient sur le caractère intégrateur des compétences «mettant en jeu des connaissances appartenant à plusieurs disciplines».

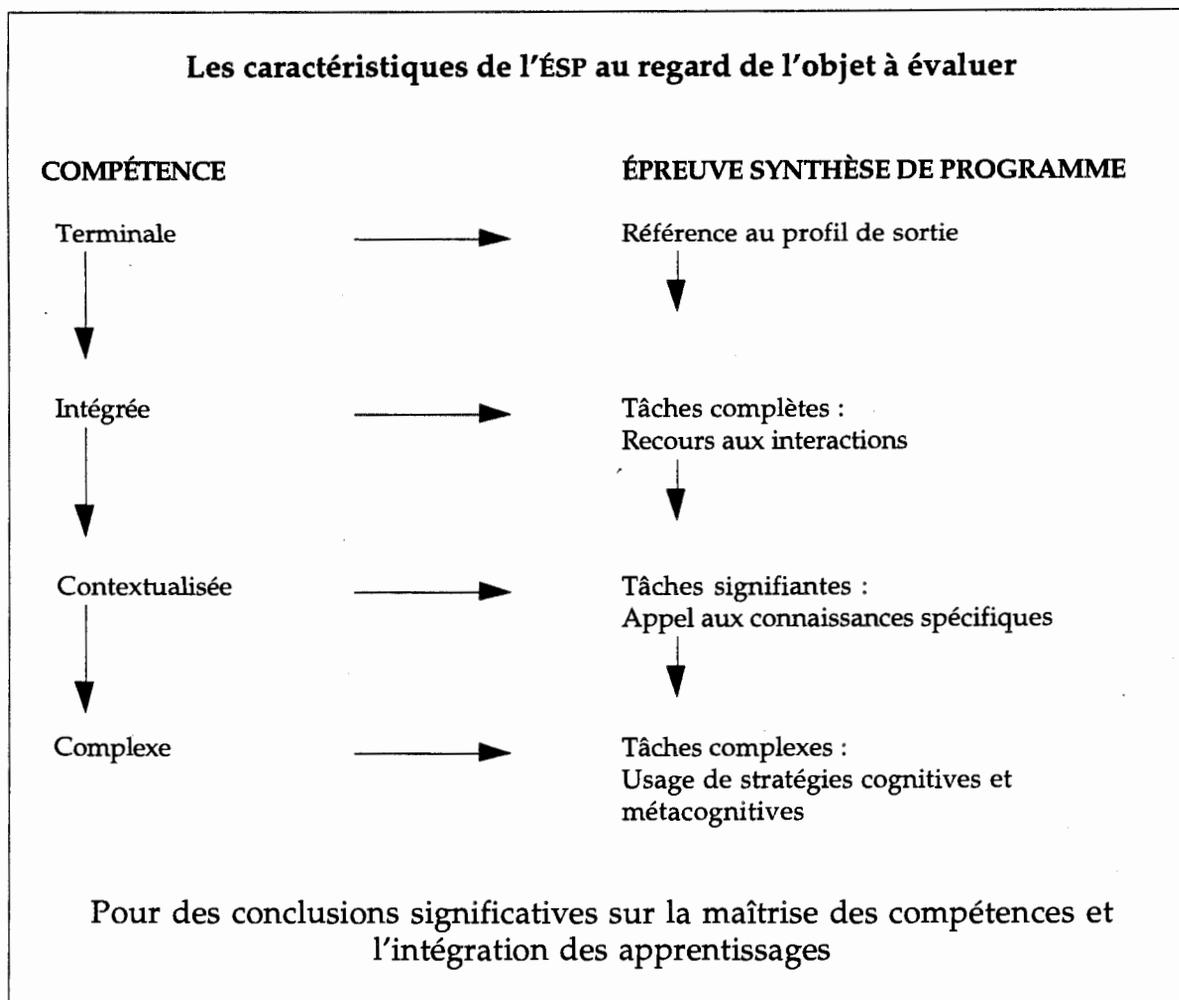
Au Collège d'Alverno dans l'État du Wisconsin, qui s'avère une source d'inspiration au Québec pour l'implantation de l'approche par compétences, on préfère le terme capacités à celui de compétences. Loacker (1995) précise que ce choix leur permet de différencier le domaine des capacités de celui des compétences qui réfère «trop souvent à des tâches compartimentées à accomplir dans des contextes spécifiques».

Pour la Direction générale de l'enseignement collégial (1994), tel que présenté dans son document «Des collègues pour le Québec du XXI^e siècle : Formation générale», les compétences et les éléments de compétence se traduisent en objectifs; les contextes de réalisations et les critères de performance tiennent lieu de standards et les activités d'apprentissage qui y sont rattachées font figure de contenus, situant ainsi le modèle dans une perspective restreinte de développement des programmes. Pour quelques programmes professionnels cependant, on a procédé à la description des situations de travail préalablement à la révision et à la rédaction des programmes en termes de compétences. Ainsi, un programme conçu en fonction de résultats attendus a plus de chance d'être en prise directe avec les situations de vie réelles.

La vision étroite des compétences que semble véhiculer le Ministère, en définissant les compétences comme de simples objectifs et sous-objectifs comportant un verbe d'action et un complément, présente l'avantage de rendre les comportements observables et donc mesurables, même si l'on confond la compétence avec une activité ou une tâche.

Dans une démarche d'élaboration d'un cadre procédural, Tremblay, pour sa part (1994b), présente les caractéristiques générales de l'épreuve synthèse au regard de l'objet à évaluer. Ces principales caractéristiques font aussi appel aux qualités essentielles de la compétence elle-même : elle correspond à un résultat terminal, elle représente un ensemble intégré, elle est contextualisée et complexe.

Le tableau suivant résume le propos de Tremblay sur les caractères particuliers de l'ÉSP déduits de ceux des compétences.



Source : Tremblay, Gilles, (1994). **L'épreuve synthèse de programme : testing ou évaluation**. Colloque de la Fédération des cégeps, Montréal.

Ces quelques variantes de l'interprétation du concept des compétences ne réussissent pas à le préciser; il demeure vague et peu opérationnel. Là encore, il faudra faire consensus sur sa définition préalablement à l'élaboration de l'ÉSP, les compétences développées dans un programme donné étant l'objet principal d'évaluation retenu par le Ministère.

2.3 Quelques nouvelles approches d'évaluation des apprentissages vs l'ÉSP

Il n'existe pas de modèle d'élaboration d'une ÉSP convenant à tous et c'est probablement un truisme de dire, comme Hager (1994) se plaît à nous le rappeler, qu'il n'existe pas de processus d'évaluation non critiquable. Des éléments importants rencontrés lors de l'examen des caractéristiques des ÉSP et de l'objet d'évaluation comme «l'utilisation des situations complexes», «l'intégration des apprentissages», les «aspects plus qualitatifs de l'évaluation», les «connaissances élargies et implicites aux compétences», les «situations de vie courante ou professionnelle» nous ont menés à des lectures plus larges sur le développement d'approches nouvelles en évaluation, notamment aux États-Unis. Ewell (1991) repère trois tendances majeures dans le domaine de la technologie de l'*assessment* : le développement de nouveaux instruments d'évaluation ayant des buts spéciaux, un changement de position important dans les pratiques évaluatives (entre autres la reconnaissance d'avenues de développement de l'étudiant qui ne sont pas linéaires) et une tendance claire vers des approches plus authentiques, plus proches du milieu naturel.

Cette incursion, bien que paraissant a priori nous éloigner du domaine de la mesure et de l'évaluation, nous a semblé nécessaire parce que le milieu collégial est très attentif à ce courant engendré par la réforme de l'éducation américaine, courant jeune et certes prometteur mais qui a son lot d'aléas et de controverses dans sa progression. Nous avons voulu relever, dans ces lectures, les aspects de ces diverses approches qui semblent convenir à la nature des ÉSP et d'autres qui paraissent incompatibles avec l'intention, la visée ou la fonction de ces épreuves synthèse.

Afin de comprendre l'engouement pour les nouvelles méthodes d'évaluation, il faut s'attarder à identifier les buts et les limites des approches traditionnelles. Le mouvement d'évaluation des compétences basé sur la Competency Based Education (CBE) et développé aux États-Unis surtout dans les années '70 utilisait, à grande échelle, les techniques du *testing* (mesures standardisées du succès des élèves). À l'origine, les tests

standardisés externes ont été développés comme un moyen d'éviter le racisme. Dans la majorité des États, on évaluait la compétence minimale des élèves de l'élémentaire et du secondaire et des étudiants des *colleges*, en lecture et écriture et en mathématiques. Le *testing* réfère souvent mais pas exclusivement aux tests papier-crayon et aux questions à choix de réponses. Il est généralement externe à la classe ou à l'institution. On lui reproche d'avoir une portée trop étroite et d'avoir des effets négatifs sur l'enseignement. L'exemple le plus courant des effets indésirables est le "*teaching to the test*" ou, dans son sens usuel, la pratique d'enseigner seulement (ou presque) les habiletés ou connaissances qui seront testées.

Certains mouvements ou courants d'évaluation ont été développés, aux États-Unis ou ailleurs, à cause d'insatisfactions grandissantes vis-à-vis les limites du *testing*. L'élargissement des méthodes d'évaluation souvent dites alternatives, terme générique englobant plusieurs approches d'évaluation, est de ces mouvements. Des appellations foisonnent dans la littérature américaine actuelle pour désigner cette forme d'évaluation plus large, plus utilitaire, en somme une alternative aux tests standardisés. Les plus courantes sont : *assessment*, *authentic assessment*, *direct assessment*, *alternative assessment*, *performance assessment* et *portfolio assessment*.

Bien que ces descripteurs reflètent des distinctions subtiles, ces nombreux types d'*assessment* présentent deux caractéristiques centrales (Worthen, 1993): premièrement, ils sont perçus comme des alternatives aux tests traditionnels standardisés constitués de questions à choix multiples; deuxièmement, tous réfèrent à l'observation directe de la performance de l'étudiant dans des tâches significatives pertinentes à la vie courante. Cette dernière particularité présente certaine similitude avec les caractéristiques de l'objet à évaluer fournis par Tremblay en ce qui a trait au contexte de la compétence. Nous y reviendrons.

Afin de bien situer le vocabulaire nouveau décrivant ces méthodes d'évaluation alternatives, McTighe et Ferrara (cités dans Worthen, 1993) mentionnent que celles-ci mettent le focus :

- sur le processus (révélé par exemple par les journaux d'apprentissage ou journaux de bord, les sessions d'observation de l'élève à qui on demande de réfléchir «tout haut», les listes de vérification en auto-évaluation);
- sur le produit (par exemple le *portfolio* d'écriture, les expositions en art);
- sur les performances (par exemple les tests de dactylo, les performances musicales, les débats oraux etc.).

L'épreuve synthèse de programme, telle que présentée actuellement au Québec, vise aussi une évaluation plus large que celle couvrant la matière d'un seul cours ou d'une seule discipline. Elle s'inspire de ce mouvement vers des méthodes alternatives, vers l'évaluation de la performance et l'évaluation des compétences. C'est pourquoi nous tenterons de donner la signification de ces termes ou de décrire brièvement quelques-unes de ces approches en signalant dans quelle mesure elles peuvent encadrer l'élaboration ou le développement des ÉSP.

2.3.1 L'*assessment*

Avant tout, il est un terme sur lequel on doit s'attarder, celui de "*l'assessment*", puisqu'il est pour ainsi dire dans tous les écrits récents anglophones et qu'il ne peut se substituer simplement au mot évaluation, en terme de jugement, comme certains sont portés à le faire. L'*assessment* est un processus de collecte d'information (Burke, 1993) pour alimenter la prise de décision et qui peut servir à divers buts à l'intérieur ou à l'extérieur de la salle de classe. L'*assessment*, longtemps associé à l'évaluation institutionnelle et à l'évaluation du personnel est de plus en plus utilisé pour l'évaluation individuelle et, dans certains cas, comme partie intégrante de l'apprentissage. Le terme *assessment* est devenu le cri de guerre du mouvement à l'encontre des problèmes reliés au *testing* et à l'évaluation des étudiants et des *colleges*, surtout depuis les années '80 (Scriven, 1991).

«L'*assessment* nécessite l'utilisation d'une variété de techniques, repose principalement sur des observations (de la performance) et exige l'intégration d'une information diversifiée dans un jugement synthèse» (Cronbach, cité dans Laliberté, 1995b). Fait à noter cependant, il ne réfère spécifiquement à aucune forme d'évaluation, sommative, formative, à interprétation normative ou critériée. Selon Engel (1994) et Simon & Gregg (1993), pour l'étudiant, l'*assessment* aide à l'apprentissage et à la mesure de ses connaissances. Pour l'enseignant, il permet de poser un diagnostic sur les apprentissages des élèves et fournit des informations permettant de prendre des décisions concernant son enseignement. Pour l'administrateur ou le public, l'*assessment* témoigne de l'efficacité d'un programme et du succès global de la formation.

Depuis les années '80, les écrits théoriques, les témoignages et les commentaires abondent dans la littérature américaine sur les approches alternatives d'*assessment* qui sont de plus en plus répandues chez nos voisins du sud, habituellement de pair avec les méthodes plus traditionnelles. Plusieurs États qui utilisent couramment ces méthodes ont, malgré tout, réservé aux tests standardisés avec questions à choix multiples les fonctions de certification ou de décision dont les enjeux sont élevés, par exemple l'attribution de diplômes ou la comparaison des résultats de différentes institutions. L'emphase mise actuellement sur les approches alternatives d'*assessment* porte sur l'information diagnostique que celui-ci permet d'obtenir et sur la possibilité d'observer et d'analyser des performances à multiples facettes plutôt que sur l'accumulation des indices globaux (comme des scores) pouvant être utilisés pour accorder des diplômes, mesurer l'efficacité des programmes ou comparer les institutions (Miller & Legg, 1993).

L'*assessment* est perçu plus efficace lorsqu'il reflète la vision que l'apprentissage est multidimensionnel, intégré et qu'il se révèle dans la performance à moyen ou long terme. L'apprentissage implique non seulement les connaissances et les habiletés mais les valeurs, attitudes et habitudes intellectuelles qui affectent et le succès académique et la performance en dehors de la classe (Peacock, 1994). L'*assessment*, comme le

révèle la racine du mot, «s'asseoir avec» (*sit with*) l'apprenant, attire l'attention sur les résultats mais aussi et également sur les expériences qui les ont générés. Pour améliorer les résultats de l'étudiant, il faut en savoir plus sur son expérience dans sa démarche d'apprentissage. L'*assessment* doit refléter cela. Pour Herman (1992), un bon *assessment* est construit sur les théories de l'apprentissage actuelles et appuyé sur les visions des habiletés et des capacités dont les étudiants auront besoin dans le futur.

Le résultat de l'ÉSP, suite à la réussite des cours d'un programme, fournira une information complémentaire permettant de porter un jugement sur les compétences de l'étudiant comme pourrait le faire (et le fait déjà dans certains programmes professionnels) un système d'*assessment* ou de suivi de la progression de l'étudiant dans les cours spécifiques à sa discipline. En effet, de nombreuses pratiques d'encadrement et d'évaluation de la performance de l'élève répandues dans plusieurs programmes du collégial supportent les décisions des enseignants : supervision directe ou indirecte de l'élève en stage en milieu de travail, rencontres formelles et régulières avec l'élève, exigence de réussite dans le volet théorique et dans le volet pratique, suivi sur place de l'élève dans ses travaux de laboratoire, progression dans les objectifs et les cours du programme ... Dans ce sens, l'*assessment* implique une démarche balisée par les mêmes visées d'une évaluation centrée sur l'intégration des apprentissages que souhaite refléter l'ÉSP. Cependant, l'*assessment* est bien une démarche de collecte de données et d'évaluation centrée autant sur le processus que sur les produits tandis que l'ÉSP est une épreuve terminale certificative donc centrée sur le produit, si l'on adhère à la vision de Cardinet (1988) mentionnée plus tôt. L'*assessment* possède une fonction diagnostique et formatrice comme telle, parfois étalée sur plusieurs mois, ce que l'ÉSP n'a pas.

2.3.2 L'*assessment* authentique

Retournons aux propos de Tremblay (chap. 2, p. 8) sur les caractéristiques des compétences, à l'élément «contextualisée». L'auteur met cette caractéristique en rapport avec le besoin de tâches signifiantes faisant appel

aux connaissances spécifiques. La tâche contextualisée nous ramène aussi à une approche nouvelle en évaluation des apprentissages, l'*authentic assessment*. Cette approche, développée en réaction aux tests à questions à choix multiples de réponses, vise une amélioration des méthodes d'enseignement et d'évaluation afin qu'elles correspondent mieux aux exigences de la vie en société et au marché du travail (Darling-Hammond, 1993). Le concept des mesures authentiques proposé il y a une dizaine d'années par Archbald & Newman (1988) domine quelques années plus tard les discussions sur l'*assessment* et a fait réagir les psychométriciens.

Pour être authentique, une tâche, ou la performance évaluée, doit être en relation avec le vrai monde ou une simulation de ce monde, c'est-à-dire une application de l'apprentissage plutôt qu'une tâche reliée purement à un contenu disciplinaire (Steele, 1992; Bergen, 1994). L'authenticité réfère donc au contexte dans lequel la performance est attendue.

D'autres caractéristiques de l'*assessment* authentique se définissent ainsi :

- Selon Archbald (1991) et Bergen (1994), il est basé sur des critères multiples: l'étudiant doit démontrer des habiletés et connaissances variées, plusieurs facettes étant évaluées à la fois. Les tâches doivent être signifiantes, complexes, plus ou moins déterminées, exigeant la sollicitation des habiletés de haut niveau et peuvent inclure la clarification des valeurs ou des aspects éthiques dans la situation donnée (Steele, 1992).
- Des standards doivent être clairement établis, écrits d'avance et être bien interprétés des juges afin de donner lieu à un jugement adéquat lors de la correction ou de l'observation (Archbald, 1991).
- L'*assessment* authentique s'appuie sur le jugement humain pour déterminer si un essai est bien écrit, si une performance est bien exécutée ou une solution bien imaginée (Archbald, 1991).

Les mots clés dans une analyse succincte de cette approche sont le contexte et le jugement. Dans un *assessment* authentique, la compétence requière les deux pour se manifester, en faisant appel aux connaissances spécifiques chez l'étudiant. Contexte et jugement sont aussi deux préoccupations centrales qui donnent de la tablature à la conception des épreuves.

L'élaboration d'épreuves dites authentiques, donc se rapprochant le plus possible du contexte réel, maximise à la fois la complexité et l'ambiguïté des tâches et la liberté de répondre de l'étudiant ce qui présente un défi majeur envers la standardisation et la fiabilité des instruments d'évaluation (Wiggins, 1993), un conflit entre la validité et la fiabilité.

La fiabilité inter-juges est aussi nécessaire dans cette approche parce que les procédures de pointage sont souvent subjectives. Plus la tâche s'avère complexe et ainsi potentiellement plus authentique, plus l'objectivité sera difficile à atteindre. Ironiquement, dans bien des cas, les situations les plus authentiques ne pourront être contraintes aux limites des épreuves ou des tests. Les évaluateurs devront expliciter davantage quelles facettes de l'authenticité sont les plus critiques (Meyer, 1992). L'entraînement des juges, la précision des standards, des critères et des listes de vérification, la clarté de la définition des tâches, la généralisabilité des résultats utilisés à des fins de sanction des études ne sont que la pointe de l'iceberg en vue d'un *assessment* authentique de qualité.

Même si une ou des épreuves authentiques semblent plus intéressantes que les méthodes traditionnelles pour reconnaître la compétence, on devra d'abord s'assurer de leur concordance avec ce qui est réellement enseigné et du fait que les étudiants aient tous une chance raisonnable de s'y préparer. Des considérations légales de taille pourraient être soulevées à cet égard comme ce fut le cas antérieurement (Jaeger, 1989).

Outre la question des coûts élevés reliés à l'évaluation des performances dans des tâches complexes authentiques (Popham, 1993, Madfaus & Kellaghan, 1993; Cusimano & al. 1994), quelques contraintes reliées aux caractéristiques de l'ÉSP risquent de refaire surface. Il a été noté par exemple

que les situations complexes ou les problèmes réels auxquels sont confrontés les étudiants dans une évaluation authentique requièrent souvent que les gens travaillent ensemble, comme un groupe, pour les résoudre (Herman, 1992), fait à ne pas perdre de vue dans l'élaboration d'une épreuve individuelle.

2.3.2.1 Le *portfolio*, un outil d'*assessment* authentique

Même s'il ne présente pas en soi une approche évaluative, nous nous attardons ici au *portfolio* puisque celui-ci a attiré une attention considérable dans la littérature et qu'il est en passe de devenir, avec la notion générale d'*assessment* des performances, la pierre angulaire de la réforme de l'éducation aux États-Unis. Un peu partout, on prône les bénéfices potentiels du *portfolio* pour l'enseignement, l'apprentissage et l'*assessment*, à nouveau en comparaison avec les tests traditionnels constitués de questions à choix multiples. Le terme *portfolio*, souvent associé à un type d'évaluation spécifique est plutôt un outil, parmi plusieurs, d'évaluation de la performance par lequel des informations peuvent être recueillies et révisées sur les travaux de l'élève dans un domaine quelconque (Reckase, 1995).

Malgré des conceptions quelque peu variées du *portfolio* dans la littérature, quelques caractéristiques y sont constamment rattachées :

- Les *portfolios* comporte des échantillons de travaux des élèves recueillis pendant un certain temps plutôt que lors d'une seule situation d'examen.
- L'évaluation du *portfolio* est basée sur des données recueillies par de multiples procédés.
- Les *portfolios* d'évaluation recueillent des échantillons de tâches régulièrement effectuées dans des contextes naturels et authentiques.
- L'évaluation du *portfolio* implique typiquement au moins deux types de données : les données brutes (échantillons des travaux des élèves) et les données résumées (compilées par le professeur). Le *portfolio* peut aussi

inclure des informations justifiant, de la part de l'élève et du professeur, le choix des productions qui y sont insérées, des brouillons de même que des pièces «satisfaisantes» et «insatisfaisantes».

- Le processus de sélection du matériel à inclure dans le *portfolio* implique, au moins jusqu'à un certain point, la participation de l'élève.

Selon Taylor (1994), la fonction principale du *portfolio* des travaux des étudiants est de recueillir un large éventail d'éléments ou de preuves pour démontrer que tous les standards, à l'intérieur d'un domaine donné, ont été rencontrés; Nolet (1992) ajoute à cela «dans un contenu ou un domaine d'habiletés que l'on ne peut observer dans un seul test».

Bien que le terme *assessment* soit lié au *portfolio* dans une variété de contextes, le *portfolio* ne sert pas exclusivement à l'évaluation de la performance de l'élève. Dans plusieurs cas, le *portfolio* sert plutôt d'outil à l'enseignement ou à l'apprentissage, contribue à faire réfléchir l'élève sur sa performance ou sa progression, le motive à lire ou écrire... Toutefois, lorsque utilisé à des fins d'évaluation, Reckase (1995) souligne que le domaine retenu doit être suffisamment bien défini de telle manière qu'un seul score soit significatif.

Les qualités que l'on semble attribuer au *portfolio* sont nombreuses (Herman & Winters, 1994; LeMahieu & al., 1995; Reckase, 1995) :

- Il représente un outil d'apprentissage important et contextualisé qui mobilise la pensée complexe et des habiletés d'expression chez l'élève.
- Il véhicule un portrait plus équitable et sensible de ce que l'élève sait ou est capable de faire.
- Il encourage les enseignants à garder le focus sur les résultats attendus de l'enseignement.
- Il fournit de l'information utile pour évaluer l'enseignement.
- Il encourage l'élève à prendre une part active dans son propre apprentissage et son *assessment* puisque celui-ci décide seul ou avec son professeur des pièces à inclure dans son *portfolio*.
- Il fournit des indices permettant d'ajuster l'enseignement.

Même si les arguments favorables à l'utilisation du *portfolio* à des fins d'enseignement sont nombreux et convaincants, on est beaucoup plus sceptiques concernant son utilisation pour des décisions ou des évaluations dont les enjeux sont élevés comme c'est le cas pour les ESP. Le *portfolio* s'avère un outil de choix en évaluation formative mais présente d'une part des problèmes sérieux de validité et de fidélité en évaluation sommative. Certains regards sceptiques portent d'autre part sur les qualités techniques et sur l'utilisation et la faisabilité du *portfolio* à large échelle. Parmi celles-ci :

- Les échantillons des travaux ou des réalisations réfèrent souvent à des domaines différents et sont obtenus dans des conditions ou des environnements variés. Ewell (1991) remarque aussi que les habiletés évaluées sont quelquefois mal définies et correspondent peu aux travaux recueillis;
- Il s'avère difficile d'obtenir un accord inter-juges satisfaisant pour un certain nombre de tâches telles que présentées dans le *portfolio*. L'accord inter-juges est un élément fondamental sur lequel s'appuient toutes les décisions sur la qualité du *portfolio*. On a aussi remarqué que les juges ont tendance à évaluer le *portfolio* sur la base d'un ou deux travaux, les meilleurs de la collection;
- Des recherches ont démontré une faible corrélation entre les résultats au *portfolio* et d'autres formes plus directes d'*assessment* (*direct assessment*) dans le même domaine (Herman & Winters, 1994; LeMahieu & al., 1995). La question à poser est alors la suivante : laquelle des évaluations représentera vraiment la capacité ou l'habileté qui persistera chez l'élève? Il y a possiblement là un problème de validité;
- Les standards appliqués pour l'évaluation de la performance sont-ils concordants avec les apprentissages à faire? avec la vision des juges «externes» (autres que les professeurs mêmes)? Les juges ont besoin de s'entendre préalablement et avec certain entraînement sur l'interprétation des critères et des standards;
- L'utilisation du *portfolio* exige beaucoup d'énergie (développement des tâches et des critères, décisions sur le choix des pièces, préparation des professeurs);

- La question d'équité ne pose pas de problèmes majeurs si le *portfolio* est utilisé comme outil d'*assessment* dans la classe puisque le professeur connaît ses élèves et a déjà d'autres indicateurs de leur compétence. Cependant, là encore, si des décisions importantes sont basées sur le *portfolio*, des interrogations réelles vont persister. Les productions sont le travail de qui? Sont-elles le produit de devoirs à la maison ou de travaux d'équipe? Les contenus des *portfolios* étant élaborés parfois sur une année entière, le support accordé à chacun des élèves dans ses travaux est-il équivalent?
- Dans certains cas, les contenus de *portfolios* de plusieurs élèves d'un même groupe peut varier beaucoup surtout si le domaine est mal défini ou les critères trop généraux. Comme le souligne Bateson (1994), les méthodes traditionnelles d'*assessment* et de *testing* basées sur la standardisation ne peuvent aider à développer et à valider des méthodes de comparaison de productions diverses comme c'est le cas dans l'élaboration du *portfolio*.

La plupart des projets ou expérimentations du *portfolio* ayant une certaine envergure et dont il est fait mention dans la littérature américaine portent sur l'évaluation ou l'*assessment* d'un seul domaine sinon d'une seule discipline, souvent les mathématiques, l'écriture, ou les arts visuels. Nous n'avons relevé, sous le vocable du *portfolio*, aucune trace de projet d'évaluation transdisciplinaire comme il est exigé des ÉSP. Malgré cela, il semble que certaines ébauches d'épreuves synthèse dans les collèges du Québec soient basées sur le *portfolio*.

Cet outil d'*assessment* présente des attraits incontestables pour les professeurs et les administrateurs avides de changement dans l'éducation. Quelques qualités ont déjà été soulignées. De surcroît, l'utilisation de cette méthode d'*assessment* dans la salle de classe motive les élèves, encourage des changements probablement très productifs dans l'enseignement, favorise la concertation entre les professeurs sur les exigences, etc. autant de renforcements souhaités dans notre réforme au collégial.

Toutefois, à notre avis, le *portfolio* en soi ne peut tenir lieu de seul outil constituant une épreuve synthèse de programme dans l'état actuel de la recherche et de notre connaissance limitée du procédé et de ses impacts, pour les raisons suivantes. Si l'on confronte les caractéristiques de l'ÉSP aux points sombres du *portfolio*, on constate que :

- Le *portfolio* est long à constituer et l'ÉSP est une épreuve terminale, donc devant tenir place à la fin du programme ou idéalement après les cours;
- Ses qualités techniques demeurant encore fragiles (fidélité, validité, équité), le *portfolio* ne peut être considéré comme un outil de choix servant à la sanction des études;
- La faisabilité du *portfolio* exige de l'entraînement des professeurs et des juges, une implantation basée sur la concertation et, dans notre cas, des programmes élaborés (et dispensés préalablement) en termes de compétences, une connaissance approfondie des différentes disciplines impliquées dans le programme, autant d'éléments peu développés dans la plupart dans nos collèges;
- Finalement, le *portfolio* élaboré même sur un seul semestre, par exemple en fin de programme, risque de poser de sérieux problèmes d'équité, le profil de l'étudiant étant susceptible d'évoluer, parfois de façon dramatique entre le début et la fin du semestre, entre le début et la fin des productions. On aura ultimement un portrait plus longitudinal que terminal de la capacité de l'élève ou de sa maîtrise des compétences.

2.3.3 Évaluation de la performance dans des tâches complexes

À la lumière des définitions et affirmations sur les compétences mentionnées au début de ce chapitre, il semble bien évident que pour évaluer les compétences des élèves au terme de leur programme, on doive, selon le domaine de l'épreuve, les soumettre à des situations problématiques suffisamment complexes devant lesquelles ils devront mobiliser leur bagages de connaissances, de capacités, d'habiletés.

Dans une perspective élargie des compétences misant sur l'intégration des apprentissages, le comportement observé, lors d'une épreuve par exemple, sera considéré comme un indicateur très imparfait, probablement difficile à mesurer. Cet indicateur permettra-t-il d'inférer le développement d'une compétence? La compétence devra-t-elle être «observée» dans une série de comportements répétés ou dans un ensemble d'actions variées dans un contexte donné simulé ou réel? L'évaluation de la performance à l'aide de tâches complètes, signifiantes, complexes remet en question toutes les pratiques traditionnelles en mesure et évaluation (Tremblay, 1994b), celles-ci étant associées depuis toujours au modèle de mesure d'un trait basé sur l'utilisation d'un grand nombre d'items compensatoires indépendants dans les tests.

L'usage croissant des mesures de la performance dans l'enseignement supérieur soulève des questions importantes à la fois sur la nature du phénomène investigué et sur l'épistémologie sous-jacente. Dans le cas des performances plus complexes, l'intention est-elle d'estimer combien d'un trait donné - ou une compétence donnée - un élève possède ou de déterminer en particulier ce qu'il peut ou ne peut faire? Dans le premier cas, l'individu devient un parmi plusieurs exemples du trait et, dans le second, le trait est inséparable de l'individu.

Dans un texte sur la validation de l'*assessment* des performances, Messick (1994) met en relief le contraste entre l'évaluation de la performance centrée sur la tâche et celle centrée sur le construit. Il distingue alors la performance et le produit à l'aide des exemples suivants. Dans des domaines comme les arts, la danse, la musique, tout ce qu'il y a à évaluer sont des performances sous une forme ou une autre. Par contre, en écriture, il y a tellement de variations acceptables dans le processus que le produit est ce qui compte principalement. Enfin dans d'autres domaines comme en mécanique ou en chimie expérimentale, et la performance et le produit garantissent la pertinence du résultat, ne serait-ce que pour éviter les accidents.

Dans le cas des performances, les questions de reproductibilité et généralisabilité des résultats ne se posent pas mais attention, cette pratique

de l'*assessment* des performances ne peut être transférée à l'évaluation des compétences ou d'autres attributs des étudiants. Dans ce cas, on ne peut inférer à partir de comportements observés que l'élève possède la connaissance ou les habiletés (des construits) sous-jacentes à ce comportement. Les compétences sont ici perçues comme un ou des construits et la performance représente un véhicule seulement et non la cible de l'évaluation. Dès lors, les critères de reproductibilité et de généralisabilité ne peuvent être ignorés et la compétence ne pourra être confirmée qu'après de nombreuses performances. Linn & al. (1991) ont aussi réservé une place de choix à la généralisabilité parmi plusieurs autres critères pour évaluer les méthodes d'*assessment* des performances. À propos du besoin de couvrir un domaine de connaissances ou d'habiletés et d'assurer une généralisabilité adéquate des résultats ils suggèrent d'augmenter, au besoin, le nombre d'épreuves pour chaque étudiant ou le nombre de tâches dans chaque épreuve. Dans le même sens, Wiggins (cité dans Brandt, 1992) précise que l'évaluation des compétences ne peut se faire sur une seule performance mais sur de nombreuses tâches ou épreuves étalées dans le temps.

2.3.3.1 La planification d'épreuves dans un système d'*assessment* basé sur les standards

Dans une tentative, fort bien réussie d'ailleurs, de faire le point sur la réforme de l'*assessment* aux États-Unis, Taylor en 1994, distingue les deux modèles d'évaluation (normative et critériée) et fournit une démarche d'élaboration d'un système d'*assessment* critérié basé sur les standards. Ce texte, qui éclaire les éducateurs et les concepteurs impliqués dans un tel processus, sans toutefois les rassurer, présente aussi une vision réaliste des risques inhérents au choix d'un système d'*assessment* nouveau. Les principaux éléments de cette démarche systématique vers l'*assessment* des standards, listés ci-après, résument l'essentiel du propos de Taylor.

En premier lieu, avant que les standards puissent être fixés, les éducateurs et autres responsables doivent s'entendre sur leurs valeurs, leurs attentes et

travailler ensemble pour identifier des buts et des résultats attendus de l'éducation qui soient importants et tangibles. À notre connaissance, cette étape pourtant fondamentale est généralement négligée sinon omise, dans nos pratiques au collégial.

La deuxième étape dans le développement des standards repose sur le choix de références ou de repères de niveau, constitués de descriptions de performances d'étudiants, à divers paliers développementaux ou étapes de formation dans la discipline et qui reflètent les résultats attendus de la formation. Parallèlement à ces repères, il faut définir les critères de performance pour chaque palier développemental, les critères étant généralement les aspects importants, la connaissance, les habiletés et le processus de pensée.

Cette tâche d'identification des critères s'avère une des plus difficiles puisque les concepteurs doivent s'entendre sur ce qui est essentiel pour reconnaître l'excellence d'une performance et aussi sur ce qui est possible à tel niveau de développement de l'étudiant. Les standards sont ensuite assortis d'exemplaires de travaux étudiants reflétant chacun des critères et des degrés de qualité pour chacun des niveaux de développement. À noter ici que les standards correspondent aux performances de qualité et non aux performances moyennes. Les standards doivent nécessairement être concordants avec les buts ou les résultats attendus et tous les exemplaires de travaux, selon les domaines ou disciplines, ne seront pas limités aux écrits mais varieront grandement selon les apprentissages à démontrer.

À propos des performances des étudiants, Taylor nous réfère à trois types :

- les tâches de performance (essais, investigations, laboratoires, présentations orales) disciplinaires ou interdisciplinaires, complétées dans le contexte de l'enseignement;
- les examens ou épreuves de performance (large éventail de tâches écrites ou de résolution de problèmes) administrées comme la plupart des examens sur lesquels sont basés des décisions;

- les *portfolios* (collections de travaux divers, écrits, résumés, projets, etc.) reliés à la discipline et élaborés sur plusieurs mois.

Chaque type d'*assessment* peut-être utilisé pour refléter l'un ou l'autre aspect des standards, des critères ou des cibles critiques des apprentissages que l'on veut évaluer (étendue des connaissances dans le domaine, habiletés de résolution de problèmes, processus, niveau de développement, organisation, usage de la langue etc.).

L'évaluation des performances dans l'atteinte des standards ne résulte pas en un score mais nécessite un jugement professionnel. Pour baliser le feedback donné à l'étudiant au besoin ou conserver de l'information sur sa progression, les juges et les enseignants ont besoin de barèmes de correction pour chacune des performances évaluées. Chaque barème doit identifier la performance et ses critères et décrire les degrés de maîtrise avec une échelle de pointage. Ce travail de description constitue la deuxième étape de standardisation des outils, la première étant celle de la description des performances de qualité dont nous avons parlé plus tôt. La performance complexe peut être jugée sur son ensemble, selon une ou plusieurs perspectives (*holistic scoring*) ou sur des dimensions séparées, spécifiques et importantes du travail (*analytic scoring*).

Une fois l'échelle des degrés de maîtrise établies, des exemplaires de travaux étudiants sont recueillis pour chaque degré de qualité décrits dans l'échelle. Ces documents de référence sont ajoutés aux standards et critères qui constituent alors des guides de correction ou d'appréciation mis à la disposition des enseignants et des élèves ou des étudiants.

La dernière étape d'élaboration d'une telle méthode d'*assessment* est probablement aussi la plus difficile à implanter. La mise à l'essai du procédé implique les enseignants et les étudiants dans la construction, l'élaboration de travaux qui démontreront clairement l'atteinte des standards. À cette étape, le professionnalisme des enseignants et la motivation des étudiants sont sollicités. Les standards et les critères étant connus de tous, les performances sont évaluées formellement, tant dans les tâches que les

portfolios seulement lorsque, de l'avis des enseignants et des étudiants, les productions et les travaux spécifiques sont à point vis-à-vis tel ou tel standard ou critère.

L'*assessment* des performances exige de repenser tout le déroulement de l'enseignement mais aussi la représentation des concepts de fidélité et de validité généralement utilisés pour évaluer les résultats ou la valeur du procédé. En grande ligne, dans un tel système, la fidélité perdra un peu de sa grande popularité au profit de la validité. La fidélité se confirmera lorsque nous aurons assez d'information recueillie sur une tâche donnée. Entretemps, il faut pratiquer et raffiner le processus de notation (Wiggins, cité dans Brandt, 1992). Quant à la validité, Linn & al. (1991) suggèrent de la scruter dans l'ensemble sous huit critères : les conséquences de l'utilisation du procédé, l'équité, le transfert et la généralisabilité des tâches, la complexité cognitive, la qualité et la couverture des domaines reliées au contenu, la signification des tâches et les coûts et l'efficacité.

2.3.3.2 L'ECOS, un exemple d'évaluation des performances complexes

Depuis une vingtaine d'années, l'examen clinique objectif structuré (ECOS) est devenu une méthode éprouvée d'évaluation de la compétence clinique des médecins et des étudiants en médecine, à la fin d'étapes longues de leur formation ou à la fin de leur programme.

Typiquement l'ECOS, développé par un médecin écossais, implique la rotation des étudiants dans un circuit de sites où on leur demande de performer dans une variété de tâches représentatives, en terme de fréquence et d'importance, de la pratique clinique usuelle du futur professionnel. Les tâches sont dites structurées parce que les habiletés sont déterminées d'avance et l'examen est dit objectif parce que tous les étudiants sont observés face à la même tâche ou problème clinique et que la correction est objective. Les examinateurs ont simplement à cocher des listes de vérification ou à donner un score global (Matte, 1994).

Ces tâches peuvent inclure une collecte d'informations auprès d'un patient, faire un examen physique, appliquer un procédé technique particulier ... L'ECOS comportant entre 10 et 20 sites permet un échantillonnage assez grand de prises de vue des habiletés cliniques, analogue à l'utilisation des questions à choix multiples de réponses pour évaluer les connaissances (Cusimano & al., 1994). Divers aspects de la performance peuvent ainsi être évalués, par exemple ceux reliés à l'attitude du candidat et ceux reliés au problème du patient en situation simulée (Cohen & al., 1991).

Bien que les qualités techniques de l'ECOS soient encore discutées (peu de fiabilité interne dans certains cas et donc généralisation difficile, peu de recherche sur la validité de construit et la validité prédictive), cette méthode présente des avantages certains par rapport aux examens à choix multiples de réponses et à l'observation directe non structurée en milieu clinique (Matte, 1994), surtout parce que le contexte se rapproche des situations réelles, est plus authentique.

Quelques stratégies permettent d'améliorer la fiabilité de l'instrument : allonger le test jusqu'à 6 heures environ, faire une sélection soigneuse des situations ayant de bonnes caractéristiques psychométriques, augmenter le nombre de sites plutôt que le nombre de juges et finalement combiner cet examen à d'autres méthodes d'évaluation, possiblement un examen écrit à réponses courtes ou à choix multiples de réponses (Newble & Swanson, 1988). Pour assurer une meilleure validité, Wiggins (cité dans Brandt, 1992) suggère de porter à six le nombre de tâches semblables pour pouvoir faire des inférences sur la maîtrise d'une tâche. Dans le cas d'épreuves dont les enjeux sont plus élevés que le feedback ou l'évaluation maison, Clauser & al. (1995) croient que huit tâches ou plus sont requises.

Cette méthode d'évaluation des performances complexes fort répandue à travers le monde des sciences de la santé (médecine, soins infirmiers, pharmacie) peut certainement servir de modèle pour l'évaluation de la performance dans le secteur professionnel des programmes collégiaux.

2.4 Par où commencer?

Cette remise en question des pratiques d'évaluation à l'aide de tâches complexes s'accompagne d'autres préoccupations fort pragmatiques devant répondre à la question «qu'allons-nous évaluer?» Les contenus des épreuves auxquelles seront soumis les élèves à l'issue d'un enseignement général ou professionnel, ne peuvent référer qu'à une seule discipline mais doivent se rapporter à un ensemble de cours faisant partie d'un programme. À titre d'exemple, une situation complexe de soins infirmiers que pourrait traiter une étudiante sortante peut exiger la sollicitation de connaissances en soins infirmiers, en biologie, en psychologie, en sociologie, la sollicitation d'habiletés de jugement clinique, de décision éthique, d'utilisation de la langue seconde, d'écriture fonctionnelle au dossier du client, de collaboration avec autrui etc. Ce champs de compétences de l'étudiante en soins infirmiers n'a pu s'agrandir adéquatement qu'avec un transfert de connaissances dans un ensemble de cours des composantes de formation spécifique, de formation générale propre ou commune voire de formation générale complémentaire de son programme.

L'apport des connaissances générales n'est certes pas négligeable dans le traitement de situations complexes et Perrenoud (1995) a souligné l'importance des connaissances dans la construction des compétences. De plus, il semble bien établi, dans les écrits sur la psychologie cognitive que l'intégration des connaissances serait favorisée par des interventions qui «insistent sur les connaissances générales dans le contexte de l'acquisition de connaissances spécifiques» et inversement que «sans ces dernières, les connaissances générales ne trouvent pas nécessairement de terrains d'application» (Tardif, 1992). La réussite de chacun des cours du programme pour obtenir le DEC permet théoriquement de vérifier l'acquisition des connaissances générales et spécifiques de l'élève.

Pour des raisons pratiques de gestion des épreuves synthèse et par méconnaissance des contenus disciplinaires ou en résistance à l'approche programme et à la description des programmes par compétences, plusieurs

intervenants du milieu collégial québécois croient que les cours de la composante de formation générale complémentaire devraient être exclus de cet ensemble global formant le programme. Lorsqu'il est question des ÉSP, le prétexte couramment évoqué est le fait que les élèves d'un même programme ne suivent pas tous les mêmes cours de cette composante.

À une autre époque et vraisemblablement dans un autre contexte, Dressel (1976) rapporte que l'examen synthèse peut porter sur la formation générale ou sur la formation spécifique seulement ou qu'il peut inclure les deux, en deux modalités d'examen, par exemple un examen oral et un examen écrit. Le contenu de l'examen peut être conçu par le département de la concentration assorti d'une quantité de lectures périphériques accessoires.

L'ÉSP doit certifier les savoirs acquis dans le contexte de la formation, confirmer la maîtrise de certaines compétences et qui mieux est, leur transférabilité dans les situations de vie courante ou professionnelle. Ces situations étant de plus en plus larges, variées et complexes en dehors du contexte et des circonstances de la formation, quelques auteurs suggèrent (St-Onge, 1995; Wiggins, 1991) d'évaluer nos étudiants en fin de programme en les confrontant à des situations de débutants sur le marché du travail ou à des standards de performance attendus chez les nouveaux étudiants à l'université.

Un certain nombre de questions se posent ici : les objectifs et les standards définis par le ministère conviennent-ils aux différents milieux de travail, sont-ils en harmonie avec les attentes plus ou moins exprimées par les institutions universitaires? Y a-t-il une distinction à faire (est-elle déjà faite?) entre des critères et standards du niveau de sortie du cégep et des critères et standards du niveau d'entrée sur le marché du travail ou aux études universitaires? Le consensus entre les employeurs et les milieux de l'enseignement et le dialogue entre les universités et les cégeps n'est pas, à notre connaissance, très éloquent sur de telles visées. Une vaste étude en soins infirmiers portant, entre autres, sur les performances idéales en situation attendues des enseignantes, des nouvelles diplômées et du

personnel soignant, révéla qu'aucun de ces groupes n'étaient d'accord sur ce que pouvait faire ou non la nouvelle diplômée (Benner, 1984).

Quoi qu'il en soit, les contenus et contextes de l'ÉSP, en nature ou en complexité, ne pourront pas dépasser ceux de la formation donnée et tenter de mesurer les compétences dans des situations qui attendent l'étudiant en milieu universitaire ou sur le marché du travail à moins d'avoir la garantie que les étudiants ont bien reçu un enseignement organisé en fonction de l'intégration des apprentissages et conséquent avec le niveau de l'épreuve. Ces situations de débutants nous apparaissent peu reproductibles dans bien des cas, surtout à la sortie des programmes préuniversitaires. À quelles situations pertinentes, valides, complexes et représentatives peut-on soumettre, par exemple, deux étudiants issus d'un même programme de sciences lettres, l'un se destinant aux études en télécommunications et l'autre en droit? Il en serait de même pour deux étudiants de sciences de la nature l'un entrant en médecine, l'autre en sciences infirmières ou en écologie. Le contenu ou l'objectif de l'évaluation sera beaucoup plus facile à cerner, on le devine, à la fin d'un programme professionnel, en autant que l'enseignement, les stages, les évaluations cours par cours et le matériel utilisé sont concordants avec les visées, les tâches du milieu du travail et, dans bien des cas, élaborés en collaboration avec ces milieux.

La mesure et l'évaluation sanctionnelle des compétences ne peut dépasser ce qui a été enseigné. Cette affirmation est lourde de conséquences puisque l'ÉSP, dans le contexte actuel, n'appartient pas à un cours mais à l'ensemble du programme. De plus, évaluer les compétences en se basant sur des standards dans un système d'évaluation à interprétation critériée ne semble pas faire partie de nos pratiques courantes d'évaluation lors de l'enseignement. À l'évidence, déterminer dans un profil de sortie ce que l'ensemble du programme développe chez l'étudiant (les compétences), sera incontournable. Un jour ou l'autre, il faudra se concerter autour des programmes, idéalement au plan national d'abord. À toutes les étapes d'élaboration, les critères de validité des épreuves seront constamment éprouvés et devront être régulièrement scrutés.

Chapitre 3

QUELQUES VISIONS ET EXPÉRIENCES D'ÉPREUVES SYNTHÈSE

Étant donné l'importance et la nouveauté de la mise en oeuvre des ÉSP dans l'enseignement collégial au Québec, il nous a semblé important de chercher dans la littérature surtout québécoise et américaine, la description théorique ou expérimentale d'épreuves de synthèse.

La troisième partie de cet essai constitue une analyse documentaire au cours de laquelle le concept d'épreuve synthèse sera abordé selon les différentes appellations, les différents niveaux d'étude ou domaines d'utilisation auxquels il réfère. Dans le cadre de cet essai, l'on s'en souviendra, l'expression épreuve synthèse renvoie à toute épreuve dont le contenu ou l'intention pédagogique dépasse un cours seulement, épreuve dont la réussite exige l'intégration des apprentissages tout au long d'un programme ou d'un ensemble de cours pouvant appartenir à des disciplines différentes.

3.1 Recherche d'information

De nombreux écrits sur les approches basées sur l'évaluation des performances ont été publiés récemment. Cependant, la plupart ne font référence qu'à une seule discipline ou fournissent des exemples quelquefois parcellaires d'évaluation des performances. Rares sont les textes décrivant des expériences d'évaluation de l'atteinte d'objectifs transdisciplinaires. Tout au mieux, quelques documents décrivent assez bien l'élaboration ou le contexte de passation d'épreuves de fin d'études dans le secteur professionnel. Plus

rare encore sont les écrits explicites d'expériences dans le secteur de formation générale.

Avant d'aborder la recherche d'information, nous nous sommes référés aux définitions que les dictionnaires ou encyclopédies nous offrent dans le sens de l'expression épreuve synthèse. Cette recherche ne nous a pas mené très loin. Legendre, dans son Dictionnaire actuel de l'éducation (1993) a défini l'examen de synthèse à partir de l'expression *comprehensive examination* selon deux sources et de la manière suivante : Activité pédagogique au cours de laquelle l'étudiant de troisième cycle doit faire preuve d'une connaissance approfondie du domaine dans lequel il se spécialise et de connaissances adéquates dans les domaines connexes (U.Q. 1988); examen dont le but est d'apprécier la maîtrise générale d'un vaste champ d'études dans le cadre de l'admission d'un candidat à des études avancées (Cepes, Unesco, 1992).

Trois éléments se dégagent de ces définitions : l'examen de synthèse réfère à la maîtrise de connaissances larges, serait surtout utilisé au niveau des études supérieures et auraient lieu à différentes étapes, à l'admission, en cours ou à la fin des études.

Pour continuer notre quête d'information, les méthodes usuelles de recherche en bibliothèque ont été utilisées et quelque information a été obtenue en localisant des documents cités dans les écrits obtenus et en considérant les citations d'un texte à l'autre.

En consultant les banques de données informatisées ERIC (Educational Research Information Center), MEDLINE, FRANCIS et PSYCHLIT, l'expression *comprehensive examination* nous a permis de repérer un certain nombre de textes portant généralement sur l'intérêt que présente depuis longtemps la notion d'épreuve synthèse au niveau des études supérieures. Afin d'investiguer l'aspect plus spécifique de vérification des compétences des étudiants avec des épreuves composées de situations complexes semblables à celles qui l'attendent à la sortie de son programme, les descripteurs suivants ont aussi été utilisés : *competency based assessment, competency evaluation, outcomes of education and achievement tests, licensing examinations,*

student educational objectives and certification, interdisciplinary instruction and tests, complex performance assessment etc., associés au descripteur *higher education*.

3.2 Classification de la documentation

Une collection d'une quinzaine de textes a été retenue pour fin d'analyse. À bien des égards, les textes ne présentent pas tous la même facture. Entre autres, les écrits issus de la communauté collégiale québécoise tels ceux du Groupe de travail sur l'épreuve synthèse de l'Université de Sherbrooke (1996), de Laliberté (1995), Munn, Lirette & Howe (1994), Tremblay (1994a et b) et St-Onge (1995) étalent évidemment des visions plus ou moins théoriques de l'ÉSP, sans expérimentation. Dans le contexte actuel, il nous a semblé intéressant de retenir ces écrits qui traitent plutôt d'un construit de l'ÉSP et qui ont introduit les responsables de l'enseignement dans le sujet. D'autres informations sont extraites de documents cadres de corporations professionnelles décrivant surtout la structure d'élaboration de leurs examens d'admission à la profession telles l'Association des infirmières et infirmiers du Canada (1993) et Les Ordres comptables agréés du Canada et des Bermudes (1995). Enfin, la plupart des textes (Anderson & al., 1984; Banta, 1986; Fong, 1988; Hager & al., 1994; Hauer, 1989; Hort & Hogan, 1988; Lenberg & Mitchell, 1991, Manus & al., 1992; Nemni, 1994) décrivent l'expérimentation d'examens à connotation de fin d'études, l'évaluation de plusieurs compétences à la fois, tant en formation générale que professionnelle et toujours au niveau des études collégiales et universitaires, jusqu'au troisième cycle d'études.

3.3 Analyse et découpage de l'information

Le tableau synoptique retrouvé en annexe, présente le sommaire de l'analyse effectuée. Les fiches sur chacun des textes contiennent le mot à mot des parties de texte représentant le mieux le thème retenu pour éviter des erreurs d'interprétation à des étapes ultérieures d'analyse.

D'autres articles sur le sujet ne comportant pas au moins 5 des 6 thèmes abordés ont été omis de ce tableau mais sont cités en référence dans le texte ou dans la bibliographie.

Le découpage du travail d'analyse en six thèmes a guidé la lecture de chacun des articles :

- 1- D'abord une **définition de l'épreuve synthèse** ou la **conception** de l'auteur dont il est question dans son texte est donnée. Les différentes définitions et conceptions regroupées dans ce tableau varient quelque peu bien que la grande majorité des auteurs croient que ces épreuves peuvent attester de l'intégration des acquis ou inférer la compétence de l'étudiant dans un domaine d'études.

D'autres imposent ces épreuves pour autoriser l'accès à une profession (AIIC, 1993; les Ordres comptables du Canada et de Bermudes, 1995), pour évaluer la qualité des programmes (Banta & Schneider, 1986) ou pour compléter un processus d'*assessment* (Fong, 1988; Hort & Hogan, 1988; Lenburg & Mitchell, 1991).

- 2- Le **domaine ou le niveau d'enseignement** est ensuite spécifié. N'ont été retenus que des textes référant à l'épreuve synthèse dans les études supérieures ou post-secondaires. Seuls les auteurs québécois (le Groupe de travail Performa, 1996; Munn, 1997; Munn, Lirette & Howe, 1994; Laliberté, 1995a; St-Onge, 1995) réfèrent à la fois aux formations générales et professionnelles ou techniques, à l'égard du niveau d'enseignement collégial et dans la perspective d'élaboration d'épreuves.

Les associations telles l'AIIC, les Ordres comptables et quelques auteurs dont Hager & al. (1994), Hort & Hogan (1988) et Lenburg & Mitchell (1991) décrivent des travaux appliqués au secteur professionnel seulement dans des disciplines spécifiques.

Quelques visions ou expérimentations d'épreuves appliquées à la formation générale sont tracées par Banta & Schneider (1986), Dressel (1976), Fong (1988), Nemni (1994). Enfin, les écrits de Anderson & al. (1984), Hauer & Slee (1989) et Manus & al. (1992) nous reportent au études de troisième cycle.

- 3- Troisièmement les **but**s de l'implantation **de l'épreuve** synthèse sont énoncés. Il est à noter que les buts de l'épreuve se distinguent parfois singulièrement de l'utilisation que l'on fait des résultats. Bien que la majorité des auteurs voient les épreuves synthèse comme l'occasion de vérifier et sanctionner l'intégration des habiletés, plusieurs y ajoutent d'autre but plus spécifique ou perçoivent l'épreuve comme un bon prétexte pour évaluer, à la fois, autre chose :
- triage des étudiants, obtention d'un grade, accès à un titre (Anderson & al., 1984; Manus & al., 1992; Hort & Hogan, 1988);
 - expérimentation d'épreuves (Fong, 1988);
 - évaluation de la capacité de communiquer (Dressel, 1976);
 - évaluation de programme (Banta & Schneider, 1986; Dressel, 1976; Munn & al., 1994);
 - protection du public (AIIC, 1993);
 - incitatif pour les étudiants à l'intégration des apprentissages (Anderson & al., 1984; Hauer & Slee, 1989; Laliberté, 1995a; Manus & al., 1992; Munn, 1997).
- 4- Sous la rubrique **description** se retrouvent, quelquefois peu détaillées, des informations qui ne sont pas du même ordre selon le texte consulté: le type d'outils ou de questions que comporte l'épreuve, la matière, le contenu ou la compétence évaluée, la longueur ou les étapes de l'épreuve. La description des épreuves fournie par la plupart des auteurs ne présente aucun modèle particulier si ce n'est Lenburg & Mitchell (1991). Certains exposent une perspective ou un cadre procédural quelquefois très général tandis que d'autres précisent la méthode ou les instruments utilisés allant jusqu'à la pondération des contenus, les types de questions etc. Les instruments utilisés ou

suggérés sont très variés, non spécifiques à un domaine ou niveau d'enseignement particulier et leur qualité est parfois discutée.

- 5- La **procédure d'élaboration ou de passation** signale, selon les auteurs, les éléments de préparation de l'épreuve par les responsables, le support ou l'encadrement offert aux étudiants dans leur préparation, quelques aspects techniques relatifs à la passation ou au calibrage de l'épreuve. L'AIRC (1993), Hort & Hogan (1988), et Lenburg & Mitchell (1991) réfèrent à une taxinomie. Dressel (1976), et Fong (1988), mentionnent l'utilisation d'examens maison ou d'épreuves achetées et quelques auteurs se prononcent sur le calibrage de leurs instruments (AIRC, 1993; Banta & Schneider, 1986; Hort & Hogan, 1988; Lenburg & Mitchell, 1991). La correction des examens et les problèmes potentiels inhérents n'est pas soulevée par plusieurs. Hort & Hogan (1988) et le Groupe de travail Performa (1996) envisagent l'utilisation d'échelles descriptives et de juges pour l'évaluation.

Le temps nécessaire à la passation des épreuves est assez souvent donné (Anderson & al., 1984; AIRC, 1993; Banta & Schneider, 1986; Dressel, 1976; Lenburg & Mitchell, 1991; les Ordres comptables du Canada et des Bermudes, 1995; Manus & al., 1992; Nemni, 1994). Un autre élément régulièrement traité est celui de la préparation nécessaire des étudiants, en terme de support, de délai, de matériel, de responsabilité (Anderson & al., 1984; AIRC, 1993; Banta & Schneider, 1986; Dressel, 1976; Hauer & Slee, 1989; les Ordres comptables du Canada et des Bermudes, 1995; Manus & al., 1992; Nemni, 1994).

Enfin, comme pour les autres thèmes, on retrouve, sous cette rubrique, une liste d'éléments à considérer dans la préparation future d'une épreuve, dans les textes qui présentent une vision théorique.

- 6- Le thème **utilisation des résultats** met en lumière, chez les auteurs qui en ont discuté, la correspondance avec le but de l'épreuve ou, sous une autre approche, l'utilisation conséquente à la disponibilité des données entourant les épreuves.

Outre la sanction des études ou d'une partie des études, utilisation la plus répandue chez la plupart des auteurs, les résultats servent, dans plusieurs cas, à d'autres fins. Banta & Schneider (1986), Dressel (1976), Fong (1988) et le Groupe de travail Performa (1996) citent l'évaluation des programmes, Hort & Hogan (1988), l'ajustement d'outils, Hager & al. (1994), le classement des étudiants, Dressel (1976), la mesure du développement des étudiants à long terme, Dressel (1976), le Groupe de travail Performa (1996), Munn (1997) et Nemni (1994), la remédiation, la vérification du progrès, le feedback à l'étudiant. Enfin, Hort & Hogan (1988) et Lenburg & Mitchell (1991) mènent, l'un une étude prédictive du succès des étudiants dans la pratique, l'autre une étude longitudinale des effets du programme d'évaluation des étudiants avec une épreuve synthèse.

Ce thème a été retenu parce que tel que mentionné plus tôt dans le chapitre 2, une épreuve ne peut servir à des fonctions très distinctes, certification, régulation, classement, évaluation de programme.

3.4 Information complémentaire

Outre les six thèmes retenus pour le découpage de l'information, la lecture de ces textes révèle des avantages et désavantages soulignés par les auteurs, de même que de nombreuses questions et réflexions suscitées par cette approche d'évaluation multidisciplinaire ou de synthèse, selon des contextes d'expérimentation ou d'élaboration de projets. Les paragraphes suivants contiennent quelques bribes d'information dans ce sens, information non comprise dans le tableau synoptique annexé mais référant toujours aux mêmes lectures.

3.4.1 Des avantages et quelques inconvénients de l'utilisation des épreuves synthèse

L'ÉSP devant implicitement mesurer l'**intégration des apprentissages**, l'utilisation d'une épreuve synthèse de fin d'étape ou fin d'études est généralement perçue, dans ces lectures, comme un facteur favorisant l'intégration des apprentissages (Laliberté, 1995a; Dressel, 1976), donnant une dimension nouvelle à l'évaluation et des conclusions socialement et professionnellement plus significatives (St-Onge, 1995; Tremblay, 1994b) permettant une évaluation des processus et des produits (Hager, & al., 1994), étant plus appropriée à la complexité ou à l'étendue des résultats à mesurer mais exigeant toutefois une adhésion aux concepts et aux méthodes utilisées de la part des professeurs, des étudiants voire des milieux, par exemple les centres de stages (Lenburg & Mitchell, 1991).

Peu d'auteurs commentent les **contenus** à évaluer dans de telles épreuves mais Hager & al., (1994) croit que d'inclure les valeurs et les attitudes est plus valide que des évaluations multiples et longitudinales sans ces aspects. L'AIIC (1993), dans son Plan directeur des examens reconnaît que l'utilisation exclusive de questions objectives à choix de réponses ne permet pas de mesurer certaines habiletés. En palliatif, on doit se fier à la validité de l'information, sur la compétence de l'étudiant, fournie par les résultats obtenus dans le programme d'études ou les cours réussis antérieurement.

À l'égard de l'**élaboration** et de l'**administration des épreuves**, à moyen terme, on évoque la lourdeur du processus. Celui-ci est généralement perçu :

- depuis toujours trop onéreux (Banta & Schneider, 1986; Dressell, 1976);
- exigeant une grande concertation dans le programme (Munn & al., 1994; Lenburg & Mitchell, 1991), beaucoup de travail sans ressources supplémentaires et devant être supporté par une structure institutionnelle (Banta & Schneider, 1986; Munn, 1997; Nemni, 1994);
- posant des problèmes d'ajustement entre les facultés et l'Ordre professionnel pour suivre l'évolution des examens, des barèmes et des

modes de correction (Les ordres comptables du Canada et des Bermudes,1995);

- impliquant des changements, des développements importants et un contrôle continu de la qualité du système et du processus (Lenburg & Mitchell, 1991);
- permettant cependant des échanges fructueux entre les départements, les institutions et les examinateurs externes (Fong, 1988).

La **qualité des épreuves** est discutée par Banta & Schneider, (1986) et l'AIC (1993), réaffirmant la crédibilité, la fiabilité et la sécurité des tests standardisés. À ce sujet, la qualité des examens élaborés par les institutions demeure discutable mais serait supérieure pour évaluer la qualité du programme. Dans ce contexte, le manque d'expertise technique compromet la validité et la fidélité de l'épreuve tandis que la sécurité exige la réécriture constante d'items. Hager & al., (1994) soulignent l'importance du contexte des épreuves pour assurer la validité tandis que Hort & Hogan (1988) trouvent plus avantageux que les standards exigés soient fixés par des membres de la profession.

Concernant la **correction des épreuves**, elle exige plusieurs juges et entraîne des problèmes de gestion de données volumineuses Hort & Hogan (1988). L'AIC et les Ordres comptables... soutiennent que les questions objectives sont plus faciles à corriger, et pour cette raison, on a augmenté le nombre de ces questions, ces récentes années, dans l'examen comptable, suite à des problèmes de complexité et de correspondance inter-juges reliées à la correction des questions construites sur la résolution de problème dans des cas complexes (Doré & Messier-Marcotte, 1994).

Autre élément non négligeable est celui des **impacts** des épreuves synthèse **chez les étudiants**. Ceux-ci acceptent plus ou moins bien la multitude de questions (Hort & Hogan, 1988), manquent souvent de préparation adéquate aux examens intégrés (Banta & Schneider, 1986) et vivent un niveau plus élevé d'anxiété lors des exercices préparatoires (Anderson & al., 1984). Malgré

ces effets indésirables, plusieurs auteurs ont noté des effets positifs, en autant que l'étudiant est informé à l'avance des objectifs à atteindre. Les étudiants, dans ce contexte :

- développent des stratégies d'étude nouvelles, plus intégrées et ont plus d'assurance et de maîtrise lors de la passation (Anderson, & al., 1984);
- savent à l'avance quoi acquérir, connaissent le niveau de performance attendu et peuvent se concentrer sur la performance désirée plutôt que sur la note (Hort & Hogan, 1988; Munn, 1997);
- sont encouragés à réviser leurs cours antérieurs dans toutes les disciplines (Dressel, 1976).

3.4.2 Des questions soulevées et des réflexions

D'entrée de jeu, certains auteurs, tels Hort & Hogan (1988), Manus & al. (1992), et St-Onge (1995), questionnent la **pertinence** globale ou celle d'un aspect des épreuves synthèse. Que font vraiment ces épreuves? Les utilise-t-on parce que tout le monde le fait? Les cas utilisés font-ils vraiment la synthèse des apprentissages? Quelle est la valeur des observations faites dans l'obtention d'un diplôme? On mentionne même que des aspects légaux pourraient être soulevés lors d'échec ou de reprise de l'épreuve dans le cas où tous les cours du programme sont préalablement tous réussis.

D'autres auteurs étalent des préoccupations d'ordre **pédagogique** : l'éducation basée sur les compétences change la relation maître-élève et dès lors, l'emphase est mise sur l'apprentissage plutôt que sur l'enseignement (Lenburg & Mitchell, 1991); l'épreuve doit être fondée sur la capacité de transfert plus que sur la rétention (D'Hainaut cité dans Laliberté, 1995a), exigence qui implique une préparation des élèves pendant tout le programme (Anderson & al., 1984). À cet égard, les Ordres comptables... déplorent que les cours soient généralement trop théoriques pendant le programme et que la formation basée sur des études de cas pratiques ne soit pas assez poussée.

Certaines **réalités des milieux** de l'enseignement confrontées aux exigences de l'ÉSP amènent des réflexions d'un autre ordre : l'implication des experts du milieu, des universités et des organismes est nécessaire, sinon souhaitée, pour assurer certaine validité des épreuves (Munn & al., 1994; AIIC, 1993); une place importante est réservée au jugement professionnel des professeurs dans l'élaboration et la correction des épreuves et, de ce fait, une formation particulière des enseignants est essentielle (Munn & al., 1994; Dressell, 1976). Enfin, le conflit entre l'ÉSP (interdisciplinaire) et l'autonomie départementale ou disciplinaire, conflit généré par l'exigence d'une approche par programme, ne peut être négligé (Dressell, 1976; Nemni, 1994). Il faut prévoir, au départ, temps et énergie pour surmonter les difficultés de consensus sur le profil de l'élève sortant, pour décider quoi mettre dans l'examen (Banta & Schneider, 1986).

Enfin, beaucoup de défis, à la fois conceptuels et pratiques, sont rencontrés dans l'élaboration des ÉSP (Munn, 1997). À l'égard de la **qualité** des épreuves, quelques auteurs émettent des recommandations voire des exigences après expérimentation ou réflexion sur le sujet : une évaluation techniquement solide présuppose un ensemble de standards de compétence valide (Hager & al., 1994). Lenburg & Mitchell (1991), dans leur vaste expérimentation, mettent l'accent sur le nombre de situations contenues dans les épreuves (parce que le praticien oeuvre dans des contextes différents) et sur la variété des méthodes à utiliser (observation directe et indirecte pour pouvoir inférer la compétence). Selon eux, l'évaluation finale ne peut être basée seulement sur la théorie ou seulement sur les épreuves dans le milieu réel. Dressell (1976) souligne réalistement, pour sa part, que les notions de contenu, de procédure, d'opportunité, d'usage des épreuves ne sont jamais résolues.

CONCLUSION

Une revue de littérature de cette sorte n'est, évidemment, jamais complète et des erreurs d'omission peuvent facilement infecter des conclusions que l'observateur tente d'en tirer. Toutefois, si des informations et quelques tendances et questions ressortent clairement de cet essai, pour permettre l'amélioration de l'enseignement et de l'évaluation en fin de programme, notre but sera atteint.

Nous avons d'abord cerné les intentions ministérielles, en concertation avec la Commission de l'évaluation, lors de l'instauration des ÉSP. Inscrite dans le Règlement sur le régime des études collégiales, l'épreuve synthèse de programme doit démontrer la synthèse des apprentissages, favoriser l'équivalence interinstitutionnelle et servir, potentiellement du moins, à l'évaluation des programmes. La mise en place de cette exigence, dans un contexte de renouveau de l'enseignement, semble discutable et onéreuse pour plusieurs.

Les nouvelles approches d'évaluation «authentiques» et généralement disciplinaires dont nous avons parlé au chapitre deux sont prometteuses à bien des égards, surtout pendant la formation suivie par l'élève plutôt qu'à la fin de son programme ou pour une évaluation ponctuelle certificative. L'évaluation terminale et sanctionnelle inhérente à l'ÉSP pose des problèmes, tant dans la conception que dans la pratique, difficultés qui ne peuvent être aplanies aisément en utilisant essentiellement l'une ou l'autre de ces approches.

Aucune des nombreuses lectures effectuées pour rédiger cet essai ne faisait état d'expérimentation d'épreuve similaire à l'ÉSP. Dans les écrits choisis, les auteurs décrivent des épreuves disciplinaires de fin d'étape ou fin d'études sans référence explicite à l'ensemble des apprentissages (par cours), dans un programme donné,

tel que suggéré dans le modèle de l'ÉSP - modèle trop peu détaillé d'ailleurs - prévu dans le renouveau.

Quelques conditions fondamentales s'imposent pour l'élaboration d'une épreuve synthèse de qualité : une connaissance approfondie du programme (ce qui n'est probablement pas le cas lors d'une réforme), un suivi systématique d'une cohorte d'élèves dans le cadre d'une approche par programme (pendant que l'organisation de l'enseignement demeure largement disciplinaire), une définition opérationnelle de l'ÉSP selon les domaines d'enseignement et un modèle d'évaluation connu de tous.

D'une part, la prescription vague de l'épreuve dans les écrits officiels et le retard dans la rédaction des programmes par compétences génèrent un handicap sérieux à l'égard de l'élaboration, la correction, l'évaluation des résultats de l'élève lors de l'ÉSP.

D'autre part, la diversité des visions quant au contenu des épreuves (référence à tous ou à une partie des cours seulement), le nombre de profils de sortie de certains programmes (en sciences humaines, foresterie, électrotechnique, sciences et techniques administratives...), le manque de concertation sur la nature des compétences, sur les standards et sur les seuils de performance, et enfin, la nécessité de dépasser le contenu d'un seul cours (même le contenu des cours d'intégration présents dans certains programmes, cours qui devraient vraisemblablement servir de préparation à l'épreuve pour l'élève) afin de créer des mises en situation réelles, nouvelles, relativement complexes lors des épreuves posent, entre autres contraintes, des défis majeurs envers les procédés méthodologiques à utiliser pour obtenir des épreuves de qualité.

Un aspect absolument unique du modèle d'évaluation selon les standards est que l'élève passe un examen lorsqu'il est prêt à le faire et qu'il accumule des performances sur une certaine période de temps plutôt que sur une seule épreuve. Actuellement, plusieurs programmes ne sont toujours pas fondés sur les compétences et standards, et même pour ceux qui le sont, à notre connaissance, l'on utilise toujours le modèle d'évaluation normative à l'intérieur de chaque cours du programme. Le maillage entre les standards utilisés pendant

l'enseignement par cours et les standards de l'épreuve synthèse, de même que la préparation des élèves à cette épreuve terminale posent un problème de validité, un problème éthique.

Est-il juste ou cohérent d'évaluer les élèves en les comparant et en les mettant en rang pendant leur programme (interprétation normative) pour ensuite les soumettre à une seule épreuve sanctionnelle dont l'évaluation sera à interprétation critériée? La comparaison de la performance de l'élève (selon les «standards» ou les seuils utilisés dans l'évaluation normative par chacun des enseignants) dans les différentes étapes ou à la fin de son programme avec sa performance à l'ESP demeurera impossible sinon problématique.

La mise en oeuvre des ESP nous impose de réviser notre façon de concevoir la réussite de l'élève, nous impose de passer d'une vision linéaire, unique, disciplinaire à une multitude de versions de la réussite ou de l'excellence. Si nous accordons de l'importance au jugement des experts et à la diversité des opinions, comme c'est le cas dans les approches d'évaluation plus authentiques et aussi dans la simulation de situations complexes, il faudra aussi revoir nos notions de fidélité, de correspondance entre les juges. Quand le modèle d'évaluation selon les standards est utilisé, la fiabilité peut être établie sur la consistance interne des performances de l'élève ce qui ne peut être mesuré lors d'une seule épreuve comportant un nombre de tâches très limité.

L'évaluation traditionnelle résume en un seul score la performance de l'élève dans un domaine donné. Dans la perspective d'utiliser des situations réelles, complexes, diversifiées à l'ESP, il deviendra crucial d'estimer la capacité de faire de l'élève, à la fois dans sa discipline ou dans son domaine mais aussi sa performance interdisciplinaire. L'on constate déjà, par exemple, que les élèves faibles en français sont nettement désavantagés dans les examens écrits à questions ouvertes. Que prévoyons-nous mesurer et évaluer à l'ESP? le français, la compétence disciplinaire ou les deux? Là encore, une question d'honnêteté des épreuves, une question éthique se pose. Pour obtenir une information utile sur l'élève dans le but de lui attribuer un diplôme, les tâches variées et complexes auxquelles il sera soumis devraient fournir des portraits différenciés de sa performance.

Il y a quelques mois, dans une lettre, la ministre acceptait le «modèle» d'épreuve synthèse proposé par les instances syndicales. Ce «modèle» intègre l'épreuve synthèse à un ou plusieurs cours en fin de programme, sous l'appellation activités synthèse de programme, si un collège souhaite le faire. La réussite de ce ou ces cours constitue la réussite de l'ÉSP. Ainsi, l'allocation des ressources nécessaires pourra-t-elle se faire via la répartition des ressources enseignantes.

Cette possibilité d'intégrer l'ÉSP aux activités courantes d'enseignement semble marquante à l'égard de l'organisation, la passation, la correction et la gestion des épreuves. Elle permet aussi, éventuellement, de vérifier la performance de l'élève sur une plus longue période de temps. Cependant, l'expression «activités synthèse de programme» laisse croire, a priori, que l'on peut noyer le poisson en cherchant des attermoissements. Dans les programmes où il n'y a pas de cours d'intégration, quel sera le cours porteur de l'ÉSP? Que deviendront le contenu et les enseignements prévus dans ce cours? Ces interrogations semblent plutôt techniques mais elles ne sont que la pointe d'un autre iceberg. Le problème demeure entier et tout aussi complexe.

En dernière analyse, la définition de l'ÉSP n'est toujours pas claire et ne le deviendra que si les responsables arrivent à se concerter sur les buts de l'éducation. La qualité des épreuves ne sera assurée que si l'on parvient à choisir et mettre en place un modèle d'évaluation applicable à l'ensemble des études. Les résultats à l'épreuve ne seront valides que si l'on prépare expressément les élèves à cet examen terminal tout au long de leurs études ou dans un cours spécifique et que l'on forme les professeurs à des pratiques d'évaluation cohérentes avec le modèle choisi.

BIBLIOGRAPHIE

- ALLAIRE, HÉLÈNE & CLAUDE MOISAN (1993). **L'évaluation des programmes de formation dans les collèges, Guide opérationnel - Deuxième version.** Commission de l'évaluation, Conseil des collèges; Québec, Gouvernement du Québec, (Collection Études et réflexions sur l'enseignement collégial), 196 p.
- ANDERSON, Wayne P., KRAUSKOPF, C.J., ROGERS, Mark E., & Gary W. NEAL (1984). Reasons for Comprehensive Examinations : A Re-evaluation. **Teaching of Psychology**, 11 (2), 78-82.
- ARCHBALD, Doug A. (1991). **Authentic Assessment : What it Means and How it Can Help Schools.** University of Wisconsin-Madison. Paper prepared for the National Center for Effective School Research and Development.
- ASSOCIATION DES INFIRMIÈRES ET INFIRMIERS DU CANADA (1993). **Plan directeur de l'examen d'autorisation infirmière à évaluation critériée.** Ottawa, AIIC, Division des examens.
- BANTA, Trudy W., & Janet A. SCHNEIDER, (1986). **Using Locally Developed Comprehensive Exams for Majors to Assess and Improve Academic Program Quality.** Paper presented at the Annual Meeting of the American Education Association, San Francisco.
- BATESON, David (1994). Psychometric and Philosophic Problems in 'Authentic' Assessment : Performance Tasks and Portfolios. **The Alberta Journal of Educational Research**, XL(2), 233-245.
- BENNER, Patricia (1984). **De novice à expert, Excellence en soins infirmiers.** Paris, Interéditions, (1995 pour la version française).
- BERGEN, Doris (1994). Authentic Performance Assessments. **Childhood Education**, (Winter '93-'94) 99-102.
- BLOOM, B. S., HASTING, J.T. & G.F. MADAUS (1971). **Handbook on Formative and Summative Evaluation of Student Learning.** New York, McGraw-Hill.

- BRANDT, RON (1992). On Performance Assessment : A Conversation with Grant Wiggins. **Educational Leadership**, May, 35-37.
- BURKE, Kay (1993). **The Mindful School : How to Assess Thoughtful Outcomes**. Palatine, Illinois, Skylight Pub.
- CARDINET, Jean (1988). **Évaluation scolaire et mesure**. Bruxelles, De Boeck-Wesmaël.
- CLAUSER, Brian E., SUBHIYAH, Raja G., NUNGESTER, Ronald G. & Douglas R. RIPKEY (1995). Scoring a Performance-Based Assessment by Modeling the Judgements of Experts. **Journal of Educational Measurement**, 32 (4), 397-415.
- COHEN, Robert, ROTHMAN, Arthur I., POLDRE, Peter & John ROSS (1991). Validity and Generalizability of Global Ratings in an Objective Structured Clinical Examination. **Academic Medecine**, 66 (9), 545-548.
- COMMISSION D'ÉVALUATION DE L'ENSEIGNEMENT COLLÉGIAL (1994). **Guide général pour les évaluations des programmes d'études réalisées par la Commission d'évaluation de l'enseignement collégial**. Québec, Gouvernement du Québec, 26 p.
- COMMISSION D'ÉVALUATION DE L'ENSEIGNEMENT COLLÉGIAL (1994). **L'évaluation des politiques institutionnelles d'évaluation des apprentissages - Cadre de référence**. Québec, Gouvernement du Québec, 20 p.
- COMMISSION D'ÉVALUATION DE L'ENSEIGNEMENT COLLÉGIAL (1994). **La Commission d'évaluation de l'enseignement collégial : sa mission et ses orientations**. Québec, Gouvernement du Québec, 21 p.
- COMMISSION D'ÉVALUATION DE L'ENSEIGNEMENT COLLÉGIAL (1995). **L'évaluation des politiques institutionnelles d'évaluation des apprentissages : premier rapport synthèse**. Québec, Gouvernement du Québec, 17 p.
- COMMISSION D'ÉVALUATION DE L'ENSEIGNEMENT COLLÉGIAL (1996). **L'évaluation des politiques institutionnelles d'évaluation des apprentissages, Rapport synthèse**. Québec, Gouvernement du Québec, 20 p.
- COMMISSION DE L'ÉVALUATION (1992). **L'examen des politiques institutionnelles d'évaluation des apprentissages et de leur mise en oeuvre**. Conseil des collèges, Québec, Gouvernement du Québec, 37 p.

- CONSEIL DES COLLÈGES (1992 avril). **L'enseignement collégial: des priorités pour un renouveau de la formation, Rapport sur l'état et les besoins de l'enseignement collégial.** Québec, Gouvernement du Québec, 413 p.
- CUSIMANO, Michael D., COHEN, Robert, TUCKER, William, MURNAGHAN, John, KODAMA, Ron & Richard REZNICK (1994). A Comparative Analysis of the Costs of Administration of an OSCE. *Academic Medicine*, 69 (7), 571-574.
- D'HAINAULT, LOUIS (1985). **Des fins aux objectifs de l'éducation, Un cadre conceptuel et une méthode générale pour établir les résultats attendus d'une formation.** Paris, Éd. Nathan.
- DARLING-HAMMOND, Linda (1994). Setting Standards for Students : The Case for Authentic Assessment. *The Educational Forum*, 59 (Fall), 14-21.
- DÉSILETS, Mario & Claude BRASSARD (1994). La notion de compétence revue et corrigée à travers la lunette cognitiviste. *Pédagogie collégiale*, 7 (4), 7-10.
- DIRECTION GÉNÉRALE DE L'ENSEIGNEMENT COLLÉGIAL (1994). **Des collèges pour le XXI^e siècle, Formation générale.** Québec, Gouvernement du Québec, 111p.
- DORÉ, Pierrette & Diane MESSIER-MARCOTTE (1994). **Analyse de la performance des candidats de l'UQAT à l'Examen uniforme de 1991 et de 1990 de l'Ordre des comptables agréés du Québec.** Rouyn-Noranda, UQAT, collection Kanasuta.
- DRESSEL, Paul L. (1976). **Handbook of Academic Evaluation**, (Chap. 11, "Comprehensive Examinations", 233-255). San Francisco, Jossey-Bass Pub.
- ENGEL, Brenda (1994). Portfolio Assessment and the New Paradigm : New Instruments and New Places. *The Educational Forum*, 29 (Fall), 22-27.
- EWELL, Peter (1991). To Capture the Ineffable : New Forms of Assessment in Higher Education. *Review of Research in Education*, 17, 75-125.
- FONG, Bobby (1988). Old Wineskins, The AAC External Examiner Project. *Liberal Education*, 74 (3), 12-16.
- FOUCAULT, CHARLES (1993). **Les épreuves synthèses : le cas des mathématiques, L'expérimentation d'un examen de service en mathématiques.** Cégep de Valleyfield. Résumé du rapport intégral par Selma Tischer, Colloque sur L'Évaluation des apprentissages. Montréal, Fédération autonome du collégial, 3p.

- GOULET, Jean-Pierre (1995). Pour ne pas en finir avec l'approche par compétence. **Pédagogie collégiale**, 8 (3), 4-7.
- GROUPE DE TRAVAIL SUR L'ÉPREUVE SYNTHÈSE (1996). **Un pas de plus vers un cadre de référence sur l'ÉSP**. Université de Sherbrooke, PERFORMA collégial.
- HAGER, Paul, Andrew GONCZI & James ATHANASOU (1994). General Issues about Assessment of Competence. **Assessment & Evaluation in higher Education**, 19 (1), 3-16.
- HAUER, Diane M., & Elisa, J. SLEE, (1989). A Review of Comprehensive Examinations in Selected Evaluation Training Programs. **Evaluation Practice**, 10 (4), 20-25.
- HERMAN, Joan L. & Lynn WINTERS (1994). Portfolio Research : A Slim Collection. **Educational Leadership**, October, 48-55.
- HORT, Linda K. & Alan E. HOGAN (1988). Setting Objectives and Assessing Competence in Professional Legal Education. **Assessment in Higher Education**, XIII (2), 92-106.
- HOWE, ROBERT & LOUISE MÉNARD (1993). **Croyances et pratiques en évaluation des apprentissages**. Laval, Collège Montmorency, (PAREA) 404 p.
- JAEGER, Richard M. (1989). Certification of Student Competence. In R.L. Linn Ed., **Educational Measurement**. New York, MacMillan.
- LALIBERTÉ, Jacques (1995a). L'épreuve synthèse de programme : gage ou facteur de l'intégration des apprentissages? **Pédagogie collégiale**, 8 (3), 18-23.
- LALIBERTÉ, Jacques (1995b). D'autres façons de concevoir et de faire l'évaluation des apprentissages. **Pédagogie collégiale**, 8 (3), 9-13.
- LEGENDRE, Renald (1993). **Dictionnaire actuel de l'éducation**. Montréal, Guérin.
- LEMAHIEU, Paul G. (1995). Portfolios in Large-Scale Assessment : Difficult But Not Impossible. **Educational Measurement : Issues and Practice**, Fall, 11-28.
- LENBURG, Carrie B & Carroll A. MITCHELL (1991). Assessment of Outcomes : The Design and Use of Real and Simulation Nursing Performance Examinations. **Nursing and Health Care**, 12 (2), 68-74.

- LES ORDRES COMPTABLES AGRÉÉS DU CANADA ET DES BERMUDES (1995). **Programme en vue de l'admission au sein de la profession de comptable agréé**, En vigueur pour l'Examen final uniforme de 1996. Toronto, Comptables agréés du Canada.
- LINN, Robert, BAKER, Eva L. & Stephen B. DUNBAR (1991). Complex, Performance-Based Assessment : Expectations and Validation Criteria. **Educational Researcher**, November, 15-21.
- LOACKER, Georgine (1995). Le développement des capacités et l'*assessment*. **Pédagogie collégiale**, 9 (1), 15-19.
- MADHAUS, George F. & Thomas KELLAGHAN (1993). The British Experience with 'Authentic' Testing. **Phi Delta Kappan**, 74(6), 458-469.
- MANUS, Michaeline B., BOWDEN, Marianne G. & E. Thomas DOWD (1992). The Purpose, Philosophy, Content, and Structure of Doctoral Comprehensive/Qualifying Exams : A Survey of Counseling Psychology Training Programs. **The Counseling Psychologist**, 20 (4), 677-688.
- MATTE, Louise (1994). **L'examen clinique objectif structuré (É.C.O.S.) comme instrument d'évaluation de la performance clinique des étudiants du programme de baccalauréat en pharmacie : un examen critique**, Essai pour l'obtention du grade de maîtrise en pédagogie des sciences de la santé. Faculté des sciences de l'éducation et Faculté de médecine, Université Laval, Québec.
- MESSICK, Samuel (1994). The Interplay of Evidence and Consequences in the Validation of Performance Assessments. **Educational Researcher**, 23(2), 13-23.
- MEYER, Carol (1992). What's the Difference Between Authentic Assessment and Performance-Based Assessment? **Educational Leadership**, May, 39-40.
- MILLER, David M. & Sue M. LEGG (1993). Alternative Assessment in a High-Stakes Environment. **Educational measurement : Issues and practices**, 12 (2), 9-15.
- MINISTÈRE DE L'ÉDUCATION (1978). **Les collèges du Québec. Nouvelle étape, Projet du gouvernement à l'endroit des Cégeps**. Québec, Gouvernement du Québec, 185 p.
- MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA SCIENCE (1993). **Des collèges pour le XXI^e siècle**. Québec, Gouvernement du Québec, 40 p.

- MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA SCIENCE (1993). **Règlement sur le régime des études collégiales**. Québec, Gouvernement du Québec, 15 p.
- MUNN, Joanne (1997). L'épreuve synthèse de programme au collégial. *Journal de l'ADMÉE*, 14(1), 11-18.
- MUNN, Joanne, LIRETTE, Nicole & Robert HOWE (1994). **Les épreuves synthèse de programme : beaucoup de questions, quelques réponses**. Québec, Colloque de l'Association québécoise de pédagogie collégiale.
- NEMNI, MAX (1994). **Examen de fin d'études, Plan du cours POL 18810**. Département des sciences politiques, Université Laval, Québec.
- NEWBLE, D. I. & D. B. SWANSON (1988). Psychometric Characteristic of the Objective Structured Clinical Examination. *Medical Education*, 22, 325-334.
- NOLET, Victor (1992). Classroom-Based Measurement and Portfolio Assessment. *Diagnostique*, 18 (1), 5-26.
- PERRENOUD, Philippe (1995). Des savoirs aux compétences, De quoi parle-t-on en parlant de compétences? *Pédagogie collégiale*, 9(1), 20-24.
- POPHAM, James (1993). Circumventing the High Costs of Authentic Assessment. *Phi Delta Kappan*, 74 (6), 470-473.
- RECKASE, Mark D. (1995). Portfolio Assessment : A Theoretical Estimate of Score Reliability. *Educational Measurement : Issues and Practice*, Spring, 12-15.
- SCRIVEN, Michael (1991). **Evaluation Thesaurus**. Newbury Park, Sage Publication, 391 p.
- SIMON, Karen & Soleil GREGG (1993). Alternative Assessment - Can Real-World Skills Be Tested? *Policy Briefs*, Apalachia Education Laboratory, West Virginia, 8p.
- SPADY, William G. (1980). The Concept and Implications of competency. In R. M. Jaeger & C. K. Tittle (Eds), **Minimum Competency Achievement Testing : Motives, Models, Measures and Consequences**. Berkeley, California. McCutchan Pub.
- ST-ONGE, Michel (1995). **Pour une épreuve synthèse de programme utile**. Cahier du renouveau no. 3, Service du développement pédagogique. Laval, Cégep Montmorency.

- STEELE, Joe M. (1992). **Performance Assessment at the College Level**. Pennsylvania Educational Research Association Annual Meeting, 8p.
- SWANSON, David B., NORMAN, Geoffrey, R. & Robert L. LINN (1995). **Performance-Based Assessment : Lessons From the Health Professions**. *Educational Researcher*, June/July, 5-11.
- TARDIF, Jacques (1992). **Pour un enseignement stratégique, L'apport de la psychologie cognitive**. Montréal, Les éditions Logiques inc.
- TAYLOR, Catherine (1994). **Assessment for Measurement or Standards : The Peril and Promise of Large-Scale Assessment Reform**. *American Educational Research Journal*, 31 (2), 231-262.
- TOURNEUR, Yvan (1985). **La certification des compétences**. *Mesure et évaluation en éducation*, 7 (4), 5-20.
- TREMBLAY, Gilles (1994a). **De l'importance d'élaborer des épreuves synthèse de programme qui soient significatives**. Québec, Colloque de l'Association québécoise de pédagogie collégiale.
- TREMBLAY, Gilles (1994b). **L'épreuve synthèse de programme : testing ou évaluation?** Montréal, Colloque de la Commission des affaires pédagogiques de la Fédération des Cégeps.
- WIGGINS, Grant (1991). **Toward One system of Education : Assessing to Improve, Not Merely Audit**. Denver, Colorado, Education Commission of the States, 38p.
- WIGGINS, Grant (1993). **Assessment : Authenticity, Context, and Validity**. *Phi Delta Kappan*, Nov, 200-214.
- WORTHEN, Blaine (1993). **Critical Issues That Will Determine the Future of Alternative Assessment**. *Phi Delta Kappan*, Feb, 444-456.

Annexe A

ÉPREUVES SYNTHÈSES : CARACTÉRISTIQUES, EXPÉRIMENTATIONS, VISIONS

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Anderson, W.P. Krauskopf, C.J. Rogers, M.E. Neal, G.W. (1984) Reasons for Comprehensive Examinations : A Re-evaluation. Teaching of Psychology, 11 (2), 78-82.</p>	<p>Philosophie : La préparation et la passation de l'examen est un pas dans le processus d'acquisition et d'application d'un ensemble d'habiletés cognitives à long terme</p>	<p>Doctorat en psychologie du counseling</p>	<ul style="list-style-type: none"> •Triage des étudiants sur la base des habiletés ou des connaissances •Rite de passage pour les étudiants (obtention du grade) •Opportunité pour les étudiants d'organiser leur pensée et d'intégrer ce qui a été appris 	<p>Questions à long développement (2 exemples donnés)</p>	<ul style="list-style-type: none"> •Examen de 10 à 12 heures •Suggestion de 4 à 6 semaines de préparation •Dossier d'anciennes questions remis aux étudiants •Suggestion de débiter l'étude directement en répondant aux questions •Feedback (dans l'étude) par les profs concernés, sur les points faibles lors des réponses aux questions 	

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Association des infirmières et infirmiers du Canada (1993) Plan directeur de l'examen d'autorisation infirmière à évaluation critériée. Ottawa, AIIC, Division des examens.</p>	<p>Examen d'autorisation infirmière qui détermine si les candidates à l'exercice de la profession possèdent les compétences d'une infirmière débutante</p>	<p>Selon les organismes d'autorisation de chaque province, après avoir suivi un programme de niveau collégial ou 1^{er} cycle universitaire reconnu</p>	<p>Protection du public. Les personnes détentrices d'un permis d'exercice possèdent les connaissances et habiletés pour exercer la profession de façon sécuritaire et efficace</p>	<p>•Porte sur 139 compétences «très importantes» et 99 «importantes» •240 à 260 questions fondées sur des cas (40 à 50%) ou indépendantes (50 à 60%) •4 niveaux de difficulté des questions •Questions à choix multiples seulement (1 énoncé et 4 options) •Compétences mesurées et leur pondération : -collecte de données 9% -analyse et interprétation des données 6% -planification des soins 7% -mise en oeuvre 41% -évaluation 10% -collaboration et coordination 4% -pratique professionnelle 22%</p>	<p>•Préparation des candidates à l'aide du «Guide canadien...» qui comporte 2 cahiers de 126 questions •Taxonomie pour les questions : -domaine cognitif (connaissance et compréhension, application, réflexion critique) -domaine affectif (réception-réponse, valorisation, organisation/ caractérisation) •Utilisation de la TRI pour le calibrage des examens à l'aide de questions d'ancrage en anglais et en français •En une journée, 2 cahiers d'examen de 3,5 heures chacun •2 sessions par année (au Québec) •Possibilité de reprise 2 fois (au Québec)</p>	<p>Sanctionne l'accès à l'exercice de la profession</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Banta, T.W. Schneider, J.A. (1986) Using locally Developed Comprehensive Exams for Majors to Assess and Improve Academic Program Quality. Paper presented At the Annual Meeting of the American Education Association, San Francisco.</p>	<ul style="list-style-type: none"> •Jusqu'aux années '80, épreuve qui évalue l'habileté de l'étudiant à utiliser ses connaissances et principes de son champ d'étude dans des situations nouvelles. •Depuis '80, on croit que ces examens peuvent attester de la qualité des programmes 	<ul style="list-style-type: none"> •Historiquement surtout dans les <i>Private Liberal Arts Colleges</i> aux USA •Dans expérimentation décrite, en enseignement général et professionnel mais sans utiliser les aspects cliniques ou les stages. 	<p>Évaluer la qualité du programme offert</p>	<ul style="list-style-type: none"> •Examens oraux ou écrits portant sur tout le programme ou le champ de concentration des études •Peu de tests à choix multiples de réponses •Longueur, forme, étendue et durée variées •Essais et tests objectifs disponibles sur le marché ou tests développés par la faculté, jumelés aux tests standardisés ou pas •Examens oraux dans petits <i>colleges</i> requièrent examinateurs externes 	<p>Expérimentation dans 11 départements :</p> <ul style="list-style-type: none"> •Définition du contenu selon les compétences centrales •Élaboration de questions : <ul style="list-style-type: none"> -à choix multiples -à courte réponse -à développement (<i>essay</i>) •Révision des items avec des spécialistes •Préparation des étudiants : <ul style="list-style-type: none"> -annonce de l'examen -suggestion d'étude à une partie d'entre eux •Passation, durée de 1 à 14 heures •Analyse d'items 	<p>À des fins d'amélioration des programmes seulement</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Dressel, P. L. (1976) Handbook of Academic Evaluation, (Chap. 11, "Comprehensive Examinations", 233-255). San Francisco, Jossey-Bass Pub.</p>	<p>Mécanisme pour encourager la révision continue et l'intégration des contenus des études actuelles et antérieures chez l'étudiant (vision répandue)</p>	<p>Historiquement surtout dans les <i>Liberal Arts Colleges</i></p>	<ul style="list-style-type: none"> •Évaluer la communication et l'intégration des connaissances •Améliorer les programmes et l'enseignement 	<p>•4 types :</p> <ul style="list-style-type: none"> -examen concernant des cours ou des séquences de cours s'étendant sur plus d'une session -examen sur des composantes des standards de l'éducation générale -examen concernant un groupe de disciplines reliées -examen limité au champ de concentration d'études •Examen oral ou écrit, tests à choix multiples, projets, dans des combinaisons variées 	<ul style="list-style-type: none"> •Examens maison ou achetés aux agences •Passation après la deuxième ou la quatrième année de <i>college</i> •Durée de quelques heures à plusieurs jours (en quatrième année) •Selon les institutions, activités supplémentaires pour préparer les étudiants : séminaires de révision, période de lecture etc. 	<p>5 utilisations :</p> <ul style="list-style-type: none"> -processus de remédiation, planification, vérification du progrès chez l'étudiant; -exigence pour graduer; -évaluation de l'enseignement et du programme; -mesure du développement à long terme des étudiants; -base partielle pour accorder des grades externes.

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
Fong, B. (1988) Old Wineskins, The AAC External Examiner Project. Liberal Education , 74 (3), 12-16.	Retour aux épreuves synthèses orales et écrites avec examinateurs externes dans un processus d' <i>assessment</i>	Majeures en éducation libérale dans 18 <i>colleges</i> et universités regroupés en triades selon l'endroit et les disciplines enseignées	Expérimentation d'instruments en vue de répondre à la question : «L'épreuve synthèse orale ou écrite, administrée par des examinateurs externes répond-elle au besoin d'avoir des approches évaluatives propres aux départements des <i>colleges</i> et universités et acceptées de leurs membres?»	Une variété de formats d'épreuves sommatives selon trois caractéristiques : •orales parfois en groupe •écrites parfois avec portfolios •avec examinateurs externes	Épreuves préparées par les membres de la faculté : -en groupe de 3 disciplines -selon les buts de la majeure du programme élaborés en département -séminaires de fins de programme dans lequel les écrits des étudiants constituent des portfolios pour les examinateurs externes	<i>Assessment</i> pour améliorer l'enseignement, l'apprentissage et la qualité des programmes

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Groupe de travail sur l'épreuve synthèse de programme (1996) Un pas de plus vers un cadre de référence sur l'ÉSP. Université de Sherbrooke, PERFORMA collégial.</p>	<p>Activité d'évaluation sommative dont l'objectif est d'attester, au terme du programme d'études, l'intégration que l'élève a réalisé des apprentissages essentiels, par ce programme</p>	<p>Collégial préuniversitaire et technique</p>	<p>Sanction partielle des études colégiales (mentionné dans d'autres documents du même groupe de travail)</p>	<p>L'épreuve couvre : •toutes les dimensions de l'intégration avec importance relative qui peut varier •acquis essentiels de la formation générale spécifique non déjà évalués de façon englobante •éléments de l'ordre des conceptions •mise en relation d'éléments fondamentaux dans le domaine •habiletés transdisciplinaires •acquis de formation générale ex. jugement éthique</p>	<p>•Élaboration : -cerner objets -préciser qualités recherchées -choisir types d'indicateurs (différents indicateurs appréhendés par différents instruments et possiblement étalés dans le temps re : stabilité ou évolution de la maîtrise de la performance et regard critique de l'élève sur ses productions) -construire instruments •Passation : -plusieurs étapes de collecte de données (pas seulement une simulation ou une étude de cas menée en quelques heures) -données recueillies en fin de dernière session ou possiblement recueillies précédemment (ex. lors de stage) et utilisées comme appui •Évaluation : -jugement avec échelle descriptive de 5 à 7 niveaux - juges entraînés et en interaction -en 2 temps : vérification des exigences et évaluation pour chaque composante</p>	<p>Perspectives : -analyser résultats pour identifier les lacunes du programme ou de la formation en même temps que celle de l'élève -occasion de fournir une rétroaction substantielle à l'élève</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Hager, P. Gonczi, A. Athanasou, J. (1994) General Issues about Assessment of Competence. Assessment and Evaluation in Higher Education, 19 (1), 3-16.</p>	<p>Dans une approche intégrée d'évaluation, accumulation d'évidences de la performance qui permet d'inférer la compétence</p>	<p>Secteur professionnel exclusivement -médecine -comptabilité (enseignement coopératif) -travail social</p>	<p>Attester de la compétence</p>	<p>Théorie et pratique dans les disciplines sont alliées dans une évaluation intégrée «holistique» parce qu'elle : •est orientée vers les problèmes •est interdisciplinaire •inclut la pratique professionnelle •couvre un champ de compétence •se concentre sur des situations courantes •exige des habiletés d'analyse •réunit la théorie et la pratique</p>	<p>•Choix de méthodes directes et variées permettant d'évaluer plusieurs éléments de compétence et tous leurs critères de performance à la fois •Regroupement des méthodes «intégrées», selon les professions : -techniques de questionnement (projets, tâches assignées, problèmes, études de cas) -simulations (patients simulés, exercices situationnels, lieu de travail simulé) -tests d'habiletés (patients standards, examen clinique structuré objectif, production) -observation directe (par superviseur, stage, expérience industrielle) -évidence d'apprentissages antérieurs (portefolio, journal de bord, répondant)</p>	<p>•Prise de décision au sujet d'un individu (pas un groupe ou un programme) •Attestation, classement, admission aux études spécialisées</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Hauer, D.M. (1989) A Review of Comprehensive Examinations in Selected Evaluation Training Programs. Evaluation Practice, 10 (4), 20-25.</p>	<p>Appellations retenues : <i>major exam</i>, <i>qualifying exam</i>, <i>comprehensive exam</i></p>	<p>Programmes de doctorat en évaluation ou autre discipline comportant une concentration en évaluation, surtout au niveau du doctorat</p>	<p>Les questions d'examen fournissent à l'étudiant la possibilité d'intégrer et de synthétiser ses connaissances dans une situation d'application</p>	<ul style="list-style-type: none"> •Essais en classe •<i>Take home</i> pour 72 heures jusqu'à une semaine par question (max. 3 questions) •Examen oral peut clarifier ou compléter l'examen écrit •Questions posées sur bibliographie préparée par l'étudiant lui-même •Possibilité pour l'étudiant de choisir ou de négocier le sujet 	<ul style="list-style-type: none"> •Examen passé en fin de scolarité avant thèse •Disponible 2 fois par année ou au besoin •Étudiant peut attendre d'être prêt avant de passer l'examen •Des groupes d'étude sont initiés par les étudiants en préparation •Système succès-échec •Examen évalué par personnes qui l'ont élaboré 	<ul style="list-style-type: none"> •Toujours possibilité de reprise •Examen élimine rarement des étudiants (<i>screening</i> ayant lieu avant)

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
Hort, L.K. Hogan, A.E. (1988) Setting Objectives and Assessing Competence in Professional Legal Education. Assessment in Higher Education, XIII (2), 92-106.	<i>Assessment</i> des compétences dans un enseignement par cas-problèmes	Dans un atelier de fin d'études (formation préparatoire professionnelle) de 32 semaines, pour les étudiants sortants de la faculté de droit	Accès au titre d'avocat ou de conseiller de la Couronne	<ul style="list-style-type: none"> •Épreuves papier crayon et séquences sur vidéo -basées exclusivement sur des cas problèmes (cohérents avec l'enseignement) -simulations de cas les plus réelles possibles •Grilles d'appréciation qualitative et numérique (processus et produit) 	<ul style="list-style-type: none"> •Après spécifications claires des habiletés à développer par rapport à tous les sujets du cours, élaboration des objectifs en terme de: -tâches terminales -de conditions de réalisation -de standards de réussite (approche de Mager) •Standards déterminés pour une interprétation critériée •Élaboration de fiches d'appréciation ayant des critères pertinents et des critères essentiels •Système d'évaluation succès-échec 	<ul style="list-style-type: none"> •Projet d'étude prédictive du succès dans la pratique légale •Ajustement des outils

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
Laliberté, J. (1995a) L'épreuve synthèse de programme : gage ou facteur de l'intégration des apprentissages? Pédagogie collégiale , 8 (3), 9-13.	Activité d'évaluation ayant pour objet l'ensemble du programme	Enseignement collégial professionnel et préuniversitaire	<ul style="list-style-type: none"> •Évaluer la qualité de l'intégration des apprentissages •Sanctionner les études 	<ul style="list-style-type: none"> •Évaluation sommative à caractère formateur •Basée sur un profil de sortie du programme 	<ul style="list-style-type: none"> •Retenir les standards et objectifs essentiels de chaque discipline du programme •Dégager des concepts intégrateurs, des matrices disciplinaires, des trames conceptuelles •Préciser la nature de l'épreuve •Prévoir préparation, encadrement et évaluation des élèves, portant attention aux standards et critères utilisés dans le processus d'enseignement 	Un élément de la sanction des études parmi d'autres

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Lenburg, C.B. Mitchell, C.A. (1991) Assessment of Outcomes : The Design and Use of Real and Simulation Nursing Performance Examinations. Nursing and Health, 12 (2), 68-74.</p>	<p>Dans une philosophie de l'éducation basée sur les compétences où l'apprentissage prévaut sur l'enseignement, examens objectifs structurés selon le <i>Lenburg's Performance Assessment Model</i></p>	<p>Baccalauréat en nursing (BSN) à l'éducation des adultes</p>	<ul style="list-style-type: none"> •Évaluer les compétences cliniques des étudiantes •Sanctionner 28 crédits du programme 	<p>4 types d'épreuves sommatives à interprétation critériée :</p> <ul style="list-style-type: none"> -les situations cliniques réelles (processus nursing complet, 3 à 5 malades avec problèmes de santé différents) -les situations quasi cliniques (examen physique et histoire complète d'un malade, enseignement au client sur des soins de santé) -les simulations en laboratoire clinique (habiletés cliniques évaluées avec des mannequins, diverses stations avec un examinateur chacune) -les simulations cliniques sur vidéo (vidéo interactif, habiletés complexes dans situations cliniques inaccessibles ou imprévisibles en milieu réel) 	<ul style="list-style-type: none"> •Concept central du modèle : la compétence - examen clinique et théorique - est documentée (non présumée) •Concept étayé de 9 autres: A- Contenu <ul style="list-style-type: none"> 1. Dimension de la pratique infirmière 2. Éléments critiques (seuils de réussite de certains comportements) B- Caractéristiques des épreuves <ul style="list-style-type: none"> 3. Échantillonnage des compétences 4. Niveau d'acceptabilité (limites d'erreur) 5. Comparabilité des épisodes de <i>testing</i> C- Structure et processus <ul style="list-style-type: none"> 6. Objectivité (documents écrits détaillés) 7. Consistance dans les jugements 8. Flexibilité (dans les environnements cliniques réels) 9. Conditions systématisées •Les épreuves durent de 1 à 2,5 jours. 	<ul style="list-style-type: none"> •Sanction des études, volet compétences cliniques •Constitution d'une banque de données pour études longitudinales sur les effets à long terme de ces épreuves et du programme d'évaluation

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Les ordres comptables agréés du Canada et des Bermudes (1995) Programme en vue de l'admission au sein de la profession de comptable agréé. En vigueur pour l'Examen final uniforme de 1996. Toronto, Comptables agréés du Canada.</p>	<p>Examen final uniforme (EFU) au Canada et aux Bermudes permettant l'accès à la profession (avec un stage obligatoire)</p>	<p>À la fin du baccalauréat en sciences comptables (universités anglophones, 4 ans et francophones, 3 ans)</p>	<p>Faire la preuve que les candidats, au moment de leur admission à la profession, maîtrisent les connaissances et les compétences essentielles à l'exercice comptable de façon acceptable</p>	<p>•L'EFU comporte 4 épreuves de 100 points avec la note de passage à 240 Il porte sur •les compétences de base : -application des connaissances de base -détermination des problèmes -analyse -intégration -jugement professionnel -satisfaction des besoins des utilisateurs -communication efficace •les connaissances de base : -connaissances approfondies -connaissances fonctionnelles -notions générales contenues dans les matières au programme et dans les contextes économique, juridique et commercial</p>	<p>•Durée de 4X4 h en 4 jours consécutifs •Cours d'appoint (15 crédits) préparatoire à l'examen •Stage obligatoire préalable à l'examen (sauf au Québec) •Calculatrice, tables de valeurs et autres fournies •Accent sur l'intégration des connaissances relevant de plus d'une matière ou plus d'un sujet •3 types de questions : -une question globale (100 points) -multidisciplinaires (160 à 200 points) -unidisciplinaires (100 à 140 points) dont des questions à choix multiples •Les questions se distinguent par : -leur niveau d'intégration -leur directivité -selon la quantité d'informations superflues qu'elles contiennent</p>	<p>Certification seulement</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Manus, M.B. Bowden, M.G. Dowd, E.T. (1992) The Purpose, Philosophy Content, and Structure of Doctoral Comprehensive/Qualifying Exam: A Survey of Counseling Psychology Training Programs. The Counseling Psychologist, 20 (4), 677-688.</p>	<p>Intention : Pourvoir à l'intégration, la synthèse, l'analyse et l'application des connaissances et habiletés</p>	<p>Doctorat en psychologie du counseling (Enquête dans facultés américaines : 57 répondants)</p>	<ul style="list-style-type: none"> •Évaluer les habiletés et intégrer les connaissances (77%) •Évaluer une large base de connaissances •Rite de passage 	<p>Examen</p> <ul style="list-style-type: none"> •écrit ou oral (58%) •écrit seulement (38%) •offert 2 fois par année ou plus •constitué de 4 à 8 questions généralement •quelquefois projets ou <i>take home</i> 	<ul style="list-style-type: none"> •Durée de 2,5 à 20 heures •Examen oral devant comité après examen écrit •Évalué par comité ou par 2 à 5 lecteurs •Possibilité de 2 ou 3 reprises après remédiation (20% des cas) pour l'examen oral •Dans quelques cas, aide aux étudiants dans la préparation ex. anciennes questions d'examen 	

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Munn, J. (1997). L'épreuve synthèse de programme au collégial. Journal de l'ADMÉE, 14 (1), 11-18.</p>	<p>Plus qu'une activité d'évaluation, acte pédagogique appuyé sur principes directeurs :</p> <ul style="list-style-type: none"> - a du sens pour l'élève - lui permet de faire le point - est explicite, équitable, a un caractère évolutif - évaluation individualisée - vérifie atteinte des compétences - est adaptée au programme et en fait partie intégrante - comprend instrumentation fidèle et valide 	<p>Enseignement collégial, en génie chimique (à titre d'exemple)</p>	<ul style="list-style-type: none"> • Juger de l'atteinte des objectifs ou des compétences fondamentales constituant le profil de sortie d'un programme • Juger de la capacité de l'élève de faire l'intégration des apprentissages réalisés et de les utiliser dans des contextes variés semblables aux situations du milieu du travail, des études universitaires ou de la vie personnelle et sociale. 	<ul style="list-style-type: none"> • Activités d'évaluation qui permettent d'observer des performances : <ul style="list-style-type: none"> - complexes - nouvelles - appellent aux habiletés intellectuelles supérieures ex. résolution de problème - reproduisent situations réelles, authentiques - ont lieu en fin de programme ou au cours de la dernière session • Activités sous forme de : <ul style="list-style-type: none"> - observation directe en laboratoire ou en stage - réalisation d'un projet ou d'une recherche - productions artistiques - communication orale - étude de cas - exmen, rapport écrit - simulation ... 	<ul style="list-style-type: none"> • Élaboration d'un profil de sortie (éléments essentiels et incontournables du programme) comportant : <ul style="list-style-type: none"> - les compétences - les types de savoirs (habiletés intellectuelles ou psychomotrices, connaissances, attitudes) • Choix des compétences complexes • Préparation des activités d'évaluation : <ul style="list-style-type: none"> - fiche descriptive (compétences, tâches de l'élève, directives - élève et correcteur, critères de correction, seuil de réussite) - instrumentation - appréciation des qualités (fidélité, validité) • Évaluation par comité 	<ul style="list-style-type: none"> • Fiche synthèse remise à l'élève ; <ul style="list-style-type: none"> - pondération et résultats - verdict - points forts et faibles • Instrumentation colligée et consignée dans portefeuille de type bilan • Appréciation d'ensemble sur des apprentissages

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Munn, J. Lirette, N. Howe, R. (1994) Les épreuves synthèses de programme : beaucoup de questions, quelques réponses, Québec, Colloque de l'AQPC.</p>	<p>Épreuve de fin de programme qui vérifie l'atteinte par les étudiants, de l'ensemble des objectifs et des standards d'un programme</p>	<p>Enseignement technique et préuniversitaire au collégial</p>	<ul style="list-style-type: none"> • Sanctionner les études • Éventuellement questionner le programme 	<p>Épreuve</p> <ul style="list-style-type: none"> • individuelle et plurielle • valide, fidèle et faisable • écrite ou orale, pratique, à longue échéance, ou combinaison de ...• 	<ul style="list-style-type: none"> • Déterminer les apprentissages à mesurer (tableau de spécification) • Choisir et élaborer instrumentation • Déterminer : <ul style="list-style-type: none"> -seuils de réussite -moments et conditions d'administration -conditions de correction -choix sur conservation et consignation des résultats • Prévoir reprise 	<ul style="list-style-type: none"> • Condition de sanction des études • Utiles pour déterminer : <ul style="list-style-type: none"> -l'inventaire des possibilités en cas d'échec -les règles pour supporter les décisions dans les cas problèmes

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Nemni, M. (1994) Examen de fin d'études, Plan du cours POL 18810. Département des sciences politiques, Université Laval, Québec.</p>	<p>Épreuve sanctionnelle en 2 volets, incluse dans un cours de fin de programme (dans cadre de réforme du programme)</p>	<p>Universitaire, premier cycle en sciences politiques</p>	<p>Cours visant à évaluer la compétence générale des étudiantes au terme des études de premier cycle</p>	<p>•Épreuve écrite de 3 questions générales, l'étudiant en choisit une •Ultérieurement, dissertation orale</p>	<p>•Épreuve écrite=60% -sans livres ni notes -en salle de classe -durée de 4 heures •Épreuve orale= 40% -devant membres du corps professoral -durée maximale de 30 minutes •Préparation des étudiants: -connaissent les 6 sujets susceptibles de faire l'objet des questions -6 rencontres facultatives sur les 6 sujets, avec conférenciers -reçoivent les questions posées lors des 8 dernières sessions d'examen •Évaluation basée sur : -la qualité de l'expression (argumentation, style, vocabulaire) -la connaissance de la matière (maîtrise du sujet, auteurs, réflexion) -la créativité (originalité, ampleur de l'analyse, critique)</p>	<p>•Sanctionne la réussite d'un cours •Permet aux étudiants de faire le bilan de leurs apprentissages •Permet aux professeurs de constater le produit de leurs efforts</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>St-Onge, M. (1995) Pour une épreuve synthèse de programme utile. Cahier du renouveau no.3, Service du développement pédagogique. Laval, Cégep Montmorency.</p>	<p>Épreuve de fin d'études qui dépasse le cumul des apprentissages tels que présentés en cours d'apprentissage</p>	<p>Au collégial •enseignement technique : toutes phases de résolution de problème associées au situations de travail •enseignement préuniversitaire : première phases de la résolution de problèmes</p>	<p>Vérifier le développement de la compétence : •capacité de faire face à des situations complexes •atteinte des standards établis par le Ministère</p>	<p>•Située au terme de la formation •Vérifie le développement des compétences (formation spécifique et générale) •Comporte des situations problématiques représentatives des débutants (marché du travail ou université)</p>	<p>Propose à l'élève une tâche •relativement complexe •nouvelle •représentative de ce qui attend le diplômé •suffisamment problématique pour être valide</p>	<p>Sanction des études</p>

Auteur et référence	Définition ou conception	Domaine ou niveau d'enseignement	But de l'épreuve	Description	Procédure d'élaboration ou de passation	Utilisation des résultats
<p>Tremblay, G. (1994a) De l'importance d'élaborer des épreuves synthèses de programme qui soient significatives. Québec, Colloque de l'AQPC. (1994b) L'épreuve synthèse de programme : testing ou évaluation? Montréal, Colloque de la Fédération des cégeps.</p>	<p>Épreuve selon conception</p> <ul style="list-style-type: none"> •systemique, •intégrée et cohérente •pertinente (valide) 	<p>Collégial</p>	<p>Avoir des conclusions significatives sur la maîtrise des compétences et l'intégration des apprentissages en fin de programme</p>	<ul style="list-style-type: none"> •Selon les caractéristiques des compétences développées dans le programme, épreuve : -terminale -intégrée -contextualisée -complexe •Épreuve sommative critériée 	<ul style="list-style-type: none"> •Cadre procédural plutôt que procédure : -référence au profil de sortie -tâches ayant recours aux interactions -appel aux connaissances spécifiques -usage de stratégies cognitives et métacognitives •Déroulement de l'épreuve en plusieurs étapes •Productions diversifiées et personnalisées. 	<p>Sanction des études dans une pédagogie de la réussite : pas de note chiffrée mais cote descriptive du degré de maîtrise</p>