

Actes du 16^e colloque de l'AQPC



MOI, J'ENSEIGNE AU COLLÉGIAL...
LE CONTEXTE ACTUEL ET SES EXIGENCES

6D58

La préparation d'une épreuve synthèse

Dominique Morissette
Professeur
Département des sciences de l'éducation
Université du Québec à Trois-Rivières



Association québécoise
de pédagogie collégiale

aqpc

LA PRÉPARATION D'UNE ÉPREUVE SYNTHÈSE

Dominique Morissette
 Département des sciences de l'éducation
 Université du Québec à Trois-Rivières

Le présent exposé est le résultat d'une réflexion personnelle concernant la pratique de l'*épreuve synthèse de programme* (ÉSP) dont on parle présentement dans plusieurs collèges de la Province. Il constitue une réaction à plusieurs articles de la revue *Pédagogie collégiale* qui traitent de ce sujet. Le point de vue adopté se fonde d'abord sur les notions de base en mesure et en évaluation pour en déduire les pratiques que nous croyons les plus souhaitables. Après avoir défini le contexte où se situe l'ÉSP et précisé la nature de l'évaluation pédagogique dont elle est une composante, nous examinons les étapes des démarches d'évaluation et de mesure, ce qui nous permet de mieux comprendre la nature et les caractéristiques de l'ÉSP. Les trois étapes de la démarche générale de mesure sont appliquées à l'enseignement où les notions d'objectifs général et spécifique puis celles de compétence et d'élément de compétence sont mises à profit. Nous considérons enfin une interprétation plausible de la qualité de justice de l'ÉSP, et les contributions de la pertinence, de la fidélité et de la validité.

Le contexte de l'épreuve synthèse

Dans la préface de l'ouvrage de Scallon (1988), Jean Cardinet présente différents contextes de l'évaluation pédagogique. Il est question de l'évaluation qui se poursuit tout au long de l'apprentissage et qui vise à soutenir l'élève dans sa progression; autrement dit, il s'agit d'un contexte formatif d'évaluation. En situation idéale, cette aide s'accompagne de rétroactions à ce point nombreuses et adaptées, qu'elle s'apparente à une forme de tutorat: chaque élève reçoit une information personnalisée, une évaluation formative individualisée. Par ailleurs, la préoccupation en évaluation peut porter non pas sur les contenus d'apprentissage, mais plutôt sur le traitement que l'élève fait subir à l'information au moment où il apprend: on parle alors d'autoévaluation, d'autopilotage ou même de métacognition. Enfin, il y a le contexte de la certification ou de la sanction des études, où l'enseignant et l'enseignante, ou l'institution se proposent de témoigner du niveau de compétence atteint par chaque élève, d'exprimer publiquement ce qu'ils pensent des apprentissages réalisés à la fin d'une période donnée d'études.

Dans ce contexte-ci, il n'est plus question de rétroaction auprès de l'élève pour corriger les faiblesses observées en cours d'apprentissage: celui-ci est terminé. Il faut plutôt se prononcer sur le niveau atteint par rapport aux contenus — processus ou produit — dont on souhaitait l'apprentissage et par rapport aux autres élèves ayant poursuivi les mêmes

objectifs, du même cours ou du même programme. Le témoignage alors rendu doit surtout être juste: pour chacun, il doit représenter le mieux possible le niveau où se situe l'atteinte des objectifs ou la maîtrise des compétences: ni plus, ni moins. Cette qualité de justice implique que celui ou celle qui a rendu ce témoignage, qui a donné cette note, puisse en garantir la qualité, en répondre même devant les tribunaux. Et c'est alors que le témoignage devra être expliqué, être motivé. Il faudra détailler comment et pourquoi on y est arrivé. Le juge devra comprendre pourquoi la note est juste.

Bref, le contexte de l'ÉSP est celui d'une évaluation sommative des apprentissages où un témoignage final est rendu, une évaluation pédagogique sommative...

- comparative: les élèves sont placés en rang, les uns par rapport aux autres (on parlera alors d'évaluation normative);
- certificative: l'enseignant et l'enseignante ou l'institution engagent leur crédibilité en témoignant du niveau de compétence atteint par chaque élève;
- critérielle: ce sont les compétences ou les objectifs prévus dans le plan de cours qui servent de fondement au témoignage: on certifie que chacun a appris ce qui était prévu.

L'ÉSP: un *assessment*, du *testing* ou une évaluation?

Dans les écrits de langue anglaise, le mot *assessment* est utilisé régulièrement. Il désigne une évaluation qui concerne une personne plutôt qu'une chose: son apprentissage, les caractéristiques de sa personnalité, ses compétences, etc. Par exemple, en parlant des directeurs d'écoles, on parlera d'un *Assessment Center* pour désigner un endroit où des candidats à la direction d'une école se présentent pour passer une batterie de tests et d'épreuves de toutes sortes dont le but est de déterminer qui a le plus de compétence pour remplir le poste disponible.

Quant au *testing*, il s'agit de l'«opération de mesure ou d'évaluation prenant place avant ou après une situation pédagogique» ou, plus généralement, de l'«action d'administrer un test», une pratique courante en psychologie ou en orientation. Le test, par rapport à l'examen ou à l'épreuve, est conçu avec plus de rigueur, en respectant les prescriptions susceptibles d'en assurer la pertinence, la

fidélité, la validité et l'utilité. Un fascicule accompagne toujours le test; le rédacteur ou la rédactrice y explique comment administrer le test et comment en interpréter les résultats. En classe, plus particulièrement dans notre milieu, le testing est rarement possible à cause des exigences de qualité imposées par de telles épreuves: l'enseignant ou l'enseignante manque de temps et parfois de compétence pour administrer de tels instruments. Voilà pourquoi on utilise des examens ou des épreuves dont la qualité est la meilleure possible, sans que soient exigées toutes les attentions accordées aux tests. Bref, le testing est un ensemble de techniques de mesure appliqué idéalement dans à peu près toutes les démarches d'évaluation des caractéristique d'une personne, d'une institution ou d'un programme, qui, en pratique, pose des exigences rarement toutes prises en compte en évaluation pédagogique.

En français, c'est donc l'expression *évaluation des apprentissages* qui représente le mieux cette action de l'enseignant ou de l'enseignante qui se propose de certifier les apprentissages ou les compétences de ses élèves ou de sanctionner leurs études. Dans le document *Principes d'équité relatifs aux pratiques d'évaluation des apprentissages scolaires au Canada* (1993) les spécialistes canadiens de l'évaluation définissent cette pratique «comme un processus qui consiste à recueillir et à interpréter des informations dans le but 1) d'informer l'élève et, s'il y a lieu, ses parents/tuteurs sur son apprentissage en termes de connaissances, d'habiletés, d'attitudes et de comportements, et 2) d'améliorer la qualité des décisions pédagogiques». Suivent cinq principes d'équité relatifs aux pratiques d'évaluation des apprentissages scolaires au Canada.

- «Les méthodes d'évaluation devraient être adaptées aux buts de l'évaluation et à son contexte général.
- On devrait offrir à tous les élèves suffisamment d'occasions de manifester les connaissances, les habiletés, les attitudes et les comportements qui font l'objet de l'évaluation.
- Les procédés utilisés pour juger ou noter la performance des élèves devraient être adaptés aux méthodes d'évaluation et appliquées de façon systématique.
- Les procédés pour faire la synthèse et l'interprétation des résultats devraient mener à des représentations justes et instructives de la performance d'un élève en relation avec les buts et les objectifs d'apprentissage pour la période visée.
- Les rapports d'évaluation devraient être clairs, précis et avoir une valeur pratique pour les personnes à qui ils s'adressent.»

En somme, même si l'enseignant ou l'enseignante utilise le testing ou la mesure ou l'observation directe au moment de préparer son témoignage, c'est de l'évaluation des

apprentissages du programme qu'il réalise, notamment au moyen de son ÉSP.

La définition de l'évaluation pédagogique au Québec

Nous avons adopté, au Québec, une définition particulière de l'évaluation des apprentissages qui a le grand avantage de mettre en relief les trois étapes d'une démarche. En effet, l'évaluation des apprentissages «consiste:

- à recueillir des informations et à les rendre significatives (mesure ou appréciation);
- à poser un jugement;
- à décider d'une action à entreprendre (décision) tenant compte des objectifs fixés au point de départ» (MEQ, 1984).

L'ÉSP n'est en fait que la première étape de cette démarche: elle consiste en une cueillette d'informations sur les apprentissages réalisés par les élèves à la fin de leur programme collégial. Il faudra bien sûr rendre ces informations significatives au regard des compétences dont on souhaite, au moment de la décision, certifier ou authentifier l'acquisition. Autrement dit il faudra démontrer la qualité de la mesure et la justice dans les décisions qui en découlent.

Bref, l'ÉSP est une opération de mesure, une cueillette d'information, une quantification de faits observés; il faut que le jugement suive, que les résultats soient appréciés tant dans leur rapport aux compétences visées que dans leurs qualités propres. Ce n'est que par la suite que la décision sera prise, que le témoignage sera rendu.

Une démarche en trois étapes

la mesure	⇒ le jugement	⇒ la décision
les faits	leur signification (l'appréciation) (leur valeur)	l'action (la note) (le témoignage)

L'ÉSP est une mesure: elle aussi se fait en trois étapes

Quel que soit le domaine où elle se pratique, la *mesure*, est essentiellement une forme de quantification. Elle consiste en l'attribution d'une valeur numérique à l'une ou l'autre des caractéristiques des objets ou des personnes. Par exemple, la mesure de la «durabilité» d'un ordinateur peut consister en l'attribution d'un nombre d'années au temps pendant lequel il a été utilisé de façon rentable. Trois étapes sont généralement franchies pour obtenir une mesure.

- Il faut d'abord définir, en utilisant des concepts, l'objet de mesure, déterminer en quoi consiste la caractéristique qui sera considérée: sa nature, ses limites, son degré d'extension ou de spécificité. C'est la présentation de l'objet d'évaluation en ayant recours à une verbalisation qui n'a rien de mesurable pour le moment, mais qui demeure un préalable nécessaire à la quantification. — titre d'exemple, il faut préciser ce que signifie la durabilité d'un ordinateur: la résistance physique de ses pièces externes ou internes ou la période d'efficacité des logiciels qu'il peut utiliser ou le temps pendant lequel sa vitesse de traitement répond aux exigences du milieu.
- Le mesure est, à bien des égards, une véritable chasse aux manifestations de ce qui a été défini ci-dessus: c'est la définition opérationnelle de l'objet de mesure, celle qui consiste à préciser les manifestations tangibles qui correspondent aux caractéristiques de l'objet de mesure. La résistance physique pourra se quantifier en nombre de jours ou d'heures d'utilisation réelle ou en nombre d'opérations exécutées; l'efficacité pourra se calculer en nombre d'opérations à la minutes ou à la seconde, etc. Il faut en somme choisir les manifestations les plus pertinentes, celles qui correspondent avec les définitions données. Et comme c'est une question de lien logique, il faudra souvent s'y mettre à plusieurs pour y parvenir.
- Enfin, il faut concevoir les instruments de quantification ou d'appréciation des manifestations énumérées ci-dessus: choisir les indices à observer ou les éléments des tâches accomplies en regard de chacune des manifestations retenues ou déterminer les preuves que les élèves doivent donner de leur compétence; construire les instrument d'observation et de quantification; retenir les techniques les mieux appropriées à la quantification de ces indices, organiser l'expression ou la présentation des résultats. Pour quantifier la résistance physique d'un clavier d'ordinateur, il faudra soumettre ses touches à des milliers de manipulations, varier de façon contrôlée la température ambiante, le faire tomber de différentes hauteurs, plusieurs fois, sur des corps de différentes textures, etc.; faire le décompte de ces indices en utilisant des appareils appropriés, puis en tenant compte des erreurs possibles.
- Il y a d'abord la définition des buts et des objectifs généraux d'un cours ou d'un programme, faite à partir des besoins des élèves. Trois sources servent habituellement d'inspiration: la clientèle concernée, le milieu social et les experts en la matière. Ces définitions sont raffinées grâce à la philosophie de l'éducation et à la psychologie de l'apprentissage.
- Suit la détermination des manifestations externes de l'apprentissage. Elles démultiplient les objectifs généraux formulés ci-dessus. Ce sont ce que plusieurs appellent les objectifs spécifiques: comportements qui constituent la preuve de l'apprentissage, par les élèves, de capacités ou d'habiletés internes. Leur détermination a d'autant plus d'importance qu'en plus de préparer la quantification des apprentissages, les objectifs spécifiques favorisent une meilleure adaptation des stratégies pédagogiques.
- Pour terminer, il y a l'organisation de l'observation directe ou indirecte de tâches tangibles que les élèves exécutent et que les enseignants et enseignantes comptabilisent en appliquant l'une ou l'autre des nombreuses techniques disponibles: les tests de toutes sortes, les épreuves écrites ou orales à réponse construite ou choisie, avec ou sans observation directe.

Dans le domaine scolaire, la mesure conserve ses trois étapes: chaque enseignant ou enseignante exprime sa représentation de l'apprentissage, traduit cette représentation sous forme de manifestations mesurables de l'apprentissage et enfin demande aux élèves d'exécuter des tâches simples ou complexes dont il quantifie les aspects qui lui semblent les plus en relation avec l'apprentissage escompté. La *pertinence* de la déduction ou la correspondance entre les attentes les plus générales, d'une part, et les plus spécifiques, d'autre part, constitue une qualité essentielle de cette démarche. En contre partie, il faut que la mesure permette d'induire, à partir des données colligées, l'atteinte des objectifs généraux et des buts et la satisfaction des besoins des élèves. Autrement dit, il est primordial que le *ruban de formation* ne soit coupé nulle part.

Quels changements apporte à ce ruban la notion de compétence?

À première vue, il semble que le changement de vocabulaire n'affecte que très peu la démarche de mesure. Il s'agit toujours du même *ruban de formation* sauf que quelques mots sont nouveaux.

- La mesure se fonde, en premier lieu sur une définition explicite des besoins des élèves; cette définition permet la description des compétences à enseigner et à apprendre dans un cours ou dans un programme puis les éléments de compétence qui en découlent.

Bref la mesure est une démarche en trois temps, même si elle n'est que la première étape de l'évaluation. Elle mérite d'autant plus l'attention qu'elle a un impact important sur les deux autres étapes qui complètent la démarche d'évaluation: le jugement et la décision.

Que deviennent ces trois étapes en enseignement-apprentissage?

En enseignement-apprentissage, on retrouve les trois mêmes étapes de la mesure, reliées de façon continue.

- Pour faire suite, il y a la précision des manifestations qui correspondent aux éléments de compétence formulés ci-dessus. C'est la définitions des objectifs spécifiques qui deviennent la preuve de l'apprentissage, par les élèves, de capacités ou d'habiletés internes, de connaissances déclaratives ou procédurales. D'autres parleront de critères et remplaceront, dans chaque énoncé, le verbe par le nom correspondant (par exemple, «définir les concepts» deviendra «définition des concepts»).
- Enfin, la dernière étape de la mesure consiste en l'organisation de l'observation directe ou indirecte de tâches tangibles que les élèves exécutent et que les enseignants et enseignantes comptabilisent en appliquant, comme il a déjà été dit, l'une ou l'autre des nombreuses techniques disponibles: les épreuves écrites ou orales à réponse construite ou choisie, avec ou sans observation systématique.

Encore ici, on distingue nettement les trois étapes de la mesure: il y est question de la définition conceptuelle, de la définition opérationnelle et de la conception et de l'administration des instruments d'observation.

À la pratique, on constate cependant que le nouveau vocabulaire génère des retombées bénéfiques, sinon du point de vue métrique, au moins du point de vue pédagogique. Ainsi, en enseignement professionnel, les apprentissages sont plus faciles à formuler et à circonscrire en énumérant les compétences et les éléments de compétence, d'autant plus qu'ils sont définis dans les cahiers de charge ou dans les descriptions des métiers ou des professions, en orientation, par exemple. En enseignement général, le fait d'exiger une formulation plus explicite des apprentissages en ayant recours aux notions de compétence et d'éléments de compétence oblige enseignants et enseignantes à réviser leurs plans de cours, à y préciser la signification donnée à l'apprentissage. Il devient plus facile de choisir des stratégies pédagogiques adaptées aux connaissances déclaratives ou procédurales enseignées, puis de mieux fixer les exigences dans les tâches de mesure. Non seulement les épreuves sont-elles améliorées, mais également tout le ruban de formation profite d'une plus grande clarté des attentes. L'enseignement étant mieux ciblé, il y a plus de chances que le temps soit utilisé à acquérir exactement ce qui mérite d'être acquis et que la mesure porte sur ce qui a été prévu puis appris.

Un inconvénient persiste malgré les nombreuses recherches pour l'atténuer: plusieurs enseignants et enseignantes analysent leur enseignement et l'apprentissage des élèves en ayant recours à un modèle cognitiviste qui demeure présentement un modèle incomplet où il est notamment difficile de trouver une hiérarchie d'apprentissage valide, de définir des stratégies d'enseignement bien adaptées, de lister des compétences définitivement utiles ou nécessaires, etc. Bref, des erreurs risquent d'être commises sans qu'on puisse en connaître la raison. Cette faiblesse découle d'une réaction assez fréquente dans notre milieu: nous sommes

portés à présenter comme principes ou comme pratiques établis ce qui n'est encore que du domaine de la recherche. L'utilisation et même parfois l'imposition généralisée de telles normes professionnelles créent de l'incertitude et de la confusion... et parfois de l'insatisfaction.

Comment se présentent les tâches particulières dans une ÉSP?

Que l'on opérationnalise les apprentissages en passant par les objectifs généraux et spécifiques ou par les compétences et les éléments de compétences, une chose est certaine: on aboutit dans un cas comme dans l'autre aux *tâches spécifiques* que l'élève doit accomplir comme preuve de son apprentissage. Or, à l'école, ces tâches plus ou moins complexes se présentent aux élèves de l'une ou l'autre des trois façons suivantes:

- l'épreuve ou la tâche à réponse construite orale ou écrite;
- l'épreuve ou la tâche à réponse choisie orale ou écrite;
- l'épreuve ou la tâche à réponse orale ou écrite observée directement.

Et si on voulait résumer ce qui doit se passer quant l'élève exécute ces tâches, on dirait qu'il doit à tout prix manifester sa compétence au niveau le plus près possible de la réalité de son apprentissage; il faut également que l'enseignante ou l'enseignant puisse se rendre compte de ce niveau de compétence et puisse en témoigner. En termes techniques, on dirait que les tâches demandées aux élèves doivent être *pertinentes* — il faut protéger la continuité du ruban de formation — et que les quantifications ou les représentations doivent être *fidèles* si on veut que les décisions soient plus *valides*. Expliquons un peu.

Dans un premier temps, il faut donc présenter aux élèves des tâches à accomplir qui reflètent le plus exactement possible ce qui leur a été enseigné en classe: des tâches *pertinentes*. Ainsi la tâche à accomplir comme preuve d'apprentissage ne peut faire appel à l'habileté d'intégration que si, au moment de l'apprentissage, et donc de l'enseignement, des tâches semblables ont été réalisées, que si cette compétence a d'abord été enseignée et apprise. On pourrait résumer les exigences de la pertinence en disant que la tâche de mesure correspondra à la compétence en trois points différents: un comportement semblable, des conditions similaires et un contenu dont on respecte les limites déjà fixée, fussent-elles élastiques.

Pour sa part, la *fidélité* fait référence à toute une série de prescriptions souvent techniques. Elle n'a pas moins d'importance que la pertinence parce qu'elle est plus technique: on a beau demander à l'élève la tâche la plus pertinente, celle qui correspond exactement à ce que le plan de cours prévoyait et à ce que l'élève a appris, si le résultat varie selon le correcteur ou selon les circonstances ou selon le moment où la correction est faite, l'injustice est tout aussi grande que si la question n'est pas appropriée. Il

faut donc choisir la sorte de tâche qui est non seulement pertinente par rapport au plan et à l'enseignement, mais également celle qui peut être administrée: présentation, correction et interprétation.

Une tendance actuelle consiste à choisir des tâches complexes observées directement plutôt que des épreuves de type crayon-papier. Du point de vue de la pertinence, on a sans doute raison, mais il ne faut pas oublier que ces tâches présentent, entre autres, deux difficultés supplémentaires:

- l'enseignant ou l'enseignante doit être capable de concevoir et d'utiliser des instruments qui le guident dans son observation: listes ou grilles, entre autres;
- il ou elle doit être familier avec les techniques d'observation afin d'éviter les dangers qu'on y rencontre habituellement: l'effet de halo, le biais personnel, l'erreur de logique et l'effet de séquence.

Entre ces deux exigences de pertinence et de fidélité, l'enseignant ou l'enseignante devra trouver sa voie ou la stratégie de mesure la mieux adaptée à la situation et à ses propres talents d'évaluateur. Ainsi parviendra-t-il à préparer les décisions les plus valides possible. Mais, au juste, qu'elle est-elle cette décision dans le cas des ÉSP?

En dernier ressort, que décide-t-on avec une ÉSP?

On l'a dit plus haut, l'épreuve synthèse de programme n'est qu'une partie de la première étape de l'évaluation des apprentissages, celle de la cueillette des informations, celle où la quantification permet de se représenter, sous un certain angle, les compétences acquises par l'élève. Reste qu'il faudra juger de la pertinence et de la fidélité de cette information et qu'il faudra y ajouter les autres informations recueillies d'autres sources. Celles-ci peuvent, selon les situations, inclure des données informelles ou non écrites, ou elles peuvent découler de la connaissance que l'enseignant ou l'enseignante a de l'élève, de ses habitudes de travail, de certains traits de sa personnalité, d'autres travaux ou devoirs scolaires comptabilisés ou pas, etc. Bien sûr, plus les sources sont incertaines, plus il faut user de prudence.

La décision, c'est le témoignage relatif à la compétence de l'élève, rendu par l'enseignant ou l'enseignante ou par une institution comme la commission scolaire, le ministère de l'Éducation ou le collège. Il prend des formes bien différentes: la lettre de référence, le profil ou le commentaire; généralement il s'agit d'une note et donc d'un ou de plusieurs symboles numériques, parfois alphabétiques. Ces symboles ont d'autant plus d'importance qu'ils engagent l'avenir de l'élève. Au Québec ou ailleurs, des personnes vont jusqu'à poursuivre en justice l'enseignant, l'enseignante ou l'institution qui a émis une note trop basse ou trop haute par rapport aux apprentissages.

Bref, pour évaluer l'élève, l'enseignant ou l'enseignante recueille de l'information, des faits, des données; il ou elle juge de la pertinence et de la fidélité de son information au regard de la décision à prendre, de l'induction à réaliser, de la note à accorder; puis il ou elle décide de cette note. La démarche vaut pour un module, un cours ou un programme; sauf que dans ce cas-ci, l'enseignant ou l'enseignante n'est pas seul et que le témoignage est rendu habituellement par une institution plutôt que par une personne.

Conclusion

- L'épreuve synthèse, qu'elle soit de programme, de cours ou de module, fait partie de l'évaluation sommative. Elle est certificative, critérielle et normative.
- L'ÉSP n'est pas toute l'évaluation sommative; elle n'en constitue que la première partie, celle où on cueille les informations. On devra par la suite les juger avant de décider.
- L'ÉSP évalue des apprentissages déjà réalisés. Ainsi, elle ne porte sur la capacité à intégrer que si cette habileté a été auparavant planifiée puis enseignée dans le programme.
- Si possible, l'ÉSP se fait par écrit (tâches à réponse construite ou choisie); la technique de l'observation peut être utilisée, mais seulement si on ne peut faire autrement. Elle est moins fidèle et exige une bonne formation de la part des enseignants et des enseignantes, et beaucoup de temps. Le gain obtenu par sa pertinence est souvent perdu par son manque de fidélité. En fin de compte les décisions risquent d'être moins valides.
- Une bonne évaluation des apprentissages exige des *tâches pertinentes, des cotes et des notes fidèles*; ainsi en arrive-t-on à des *décisions* qui ont de bonnes chances d'être *valides*.

Références

COMITÉ CONSULTATIF MIXTE. *Principes d'équité relatifs aux pratiques d'évaluation des apprentissages scolaires au Canada*, Centre for Research in Applied Measurement and Evaluation, Université de l'Alberta, 1993, 21 p.

MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (MEQ). *Éléments de docimologie. 2. Lexique*, Gouvernement du Québec, Québec, mars 1984, 43 p.

MORISSETTE, Dominique. *L'évaluation sommative: guide pratique*, Saint-Laurent, Éditions du Renouveau Pédagogique inc., 1996 (sous presses).

SCALLON, Gérard. *L'évaluation formative des apprentissages (Tome I: La réflexion)*, Québec, Les Presses de l'Université Laval, 1988, 171 p.