

Les pratiques évaluatives en éducation physique à l'ordre d'enseignement collégial

RAPPORT D'ANALYSE DU SONDAGE

Gilles Champagne, Luc Chiasson,
Robert Delaney et François Lasnier

En collaboration avec



Québec, février 2013

Cette recherche a été proposée par le comité d'enseignants et d'enseignantes en éducation physique coordonné par Mme Francine Lamarre du Cégep de La Pocatière et supportée par la Fédération des éducateurs et éducatrices physiques enseignants du Québec (FÉEPEQ)

Le contenu du présent rapport n'engage que la responsabilité de ses auteurs.

On peut obtenir une copie ce rapport de recherche en fichier PDF en s'adressant à la :

FÉEPEQ

Université de Sherbrooke - FEPS

2500, boul. de l'Université

Sherbrooke (Québec)

J1K 2R1

Téléphone : 819.821.8000 poste 63799

Télécopieur : 819.821.7970

info@feepq.com

<http://www.feepeq.com>



La fédération des éducateurs et éducatrices physiques enseignants du Québec (FÉÉPEQ) est heureuse de présenter les résultats d'une enquête concernant les pratiques évaluatives de professeurs en éducation physique de niveau collégial au Québec. Cette enquête vous permettra de connaître l'état de la situation au niveau des pratiques évaluatives employées par les professeurs en éducation physique dans les Cégeps. Elle vous permettra également d'identifier les besoins des professeurs en termes de formation continue en lien avec l'approche par compétence et les pratiques évaluatives en lien avec cette approche. Nous espérons avoir le privilège de vous servir de nouveau dans un avenir rapproché.

Patrick Parent
Président

COMITÉ PROVINCIAL POUR LA DISCIPLINE ÉDUCATION PHYSIQUE

L'évaluation de l'atteinte des compétences a toujours été au cœur de nos préoccupations. Le projet d'actualisation des devis était un moment propice pour effectuer un sondage sur les pratiques évaluatives en éducation physique au collégial. Cette enquête est le fruit d'un travail colossal dirigé par des passionnés de l'éducation physique au niveau collégial. Les résultats nous aideront à identifier nos besoins perfectionnements ainsi que dans nos tâches d'évaluateur. C'était un privilège pour notre comité d'être associé à cette étude. Nous en remercions ses auteurs.

Francine Lamarre, Cégep de La Pocatière

Responsable du comité provincial d'éducation physique

RÉSUMÉ

Cette étude vise à décrire les pratiques évaluatives des enseignants et enseignantes en éducation physique des cégeps du Québec afin de les aider à s'améliorer.

Un questionnaire a été développé et validé, à l'aide d'un comité d'experts, afin de mesurer les concepts ou regroupements des pratiques évaluatives suivantes : modalités d'évaluation, instruments de mesure, instruments de jugement, composition de la note et qualités des évaluations. Le questionnaire a été traduit en anglais afin de pouvoir sonder tous les enseignants et enseignantes en éducation physique du réseau collégial québécois. Le questionnaire a été expédié par courrier électronique à tous les professeurs en éducation physique de tous les cégeps francophones et anglophones. Notre échantillon est composé de 131 sujets. L'analyse des données a distingué les particularités de chacune des trois compétences en éducation physique.

Les résultats ont permis d'identifier les pratiques évaluatives les plus fréquentes et les moins fréquentes. De plus, pour chaque concept analysé, on a estimé le degré de proximité des pratiques évaluatives en fonction de la formation par objectifs (FPO) et en fonction de la formation par compétences (FPC). Pour certains concepts, les professeurs sont davantage orientés vers la FPC, et pour d'autres, ils sont plus orientés vers la FPO. Globalement, les pratiques évaluatives sont plus orientées vers la FPO que la FPC. Par contre, en général, les bonnes pratiques évaluatives sont nettement plus utilisées que les mauvaises pratiques.

Il y aurait certainement avantage à poursuivre l'étude des pratiques évaluatives en vue de mieux comprendre l'application de certaines pratiques, mais la présente étude a permis d'identifier clairement des secteurs de perfectionnement pour les enseignants et les enseignantes en éducation physique du réseau collégial.

ABSTRACT

This study aims to characterize the evaluation practices of CÉGEP Physical Education teachers in order to help them develop knowledge and awareness of, as well as skills in evaluation.

A questionnaire was developed and validated with the help of a panel of experts in order to measure the concepts or groupings of the following evaluation practices: evaluation procedures, measurement tools, judgement tools, the make-up of the grade and quality of the evaluation. The French questionnaire was translated into English and sent to the French and English CEGEPS in order to survey all CEGEP Physical Education teachers working in the Quebec College network. Our sample is composed of 131 subjects. The analysis of the data distinguished specific aspects of each of the three competencies in physical education.

The results permitted the identification of the most frequent and least frequent evaluation practices in Physical Education. Furthermore, an estimation of the strength of the relationship of these practices to objective-based learning (OBL) and competency-based learning (CBL) for each of the concepts was made. For certain concepts, the teachers are more in line with CBL, whereas for others they are more in line with OBL. Overall, the evaluation practices are more in line with OBL than with CBL. However, in general, good evaluation practices are used much more than are bad evaluation practices.

It would be advantageous to continue the study of these evaluation practices in order to better understand the application of certain practices. However, the present study did identify distinct areas in evaluation-for which professional development would benefit CEGEP Physical Education teachers.

FAITS SAILLANTS

But de l'étude

Décrire les pratiques d'évaluation des apprentissages des enseignants et des enseignantes en éducation physique des cégeps du Québec afin de les aider à s'améliorer.

Cette section présente une brève définition des concepts analysés, ainsi que les pratiques évaluatives les plus fréquentes et les moins fréquentes.

MODALITES D'ÉVALUATION

Les modalités concernent principalement les objets d'évaluation (connaissances, habiletés, compétences, éléments de compétence, produit, processus, intégration, autres). Les modalités réfèrent aussi aux méthodes d'évaluation, au contexte, aux limites et contraintes, au nombre et aux types d'évaluation. Elles portent aussi sur la *Politique institutionnelle d'évaluation des apprentissages* et les règles départementales.

Modalités d'évaluation les moins fréquentes

- Double seuil
- Tests de condition physique non standardisés
- Performances physiques en situation de jeu
- Évaluations faisant appel au par cœur
- Trois évaluations et moins par session

Modalités d'évaluation les plus fréquentes

- Évaluation de la compétence en se basant sur les éléments de la compétence et les critères de performance
- Pondération des critères selon leur importance
- Informations données aux étudiants à l'avance par rapport aux modalités d'évaluation
- Évaluations conformes aux directives du département
- Plagiat (détection et sanction)

Formation par objectifs (FPO) vs formation par compétences (FPC)

FPO forts

- Évaluations de 12 % et moins
- Plus de 4 évaluations par session
- Moins de 40 % des points pour l'évaluation finale

FPC faibles

- 3 évaluations et moins par session
- 40 % et plus des points pour l'évaluation finale

INSTRUMENTS DE MESURE

Type de tests et d'examens utilisés, types de tâches d'évaluation, modes de cueillettes d'information et application des instruments de mesure.

Instruments de mesure les moins utilisés

- Entrevues
- Questions au choix
- Exposés oraux

Instruments de mesure les plus utilisés

- Carnet de bord
- Tests de condition physique
- Évaluation de la compétence avec des tâches d'intégration
- Tâches d'intégration inscrites dans le plan de cours
- Examens théoriques écrits, composés de questions à réponses courtes
- Épreuve synthèse de cours

Formation par objectifs (FPO) vs formation par compétences (FPC)

FPO fort

- Examens théoriques écrits
- Questions à réponses courtes

FPC faible

- Le choix des instruments de mesure est bon (cahier de bord ou tâche d'intégration)
- Cependant, on doute de la capacité des instruments de mesure à bien intégrer les éléments de compétence en un tout

INSTRUMENTS DE JUGEMENT

Les instruments de jugement servent à porter un jugement à partir des instruments de mesure. Ils concernent les critères de performance, les critères d'évaluation, les éléments de la compétence, les différents types de grilles.

Instruments de jugement les moins utilisés

- Je note à partir d'une cote
- Critères autres que les critères de performance
- Sélection par soi-même des critères d'évaluation de la compétence

Instruments de jugement les plus utilisés

- Utilisation des critères d'évaluation sélectionnés par le département
- Utilisation des éléments de la compétence comme critères
- Utilisation de tous les critères de performance de la compétence

Formation par objectifs (FPO) vs formation par compétences (FPC)

FPO fort

- Évaluation de tous les critères de performance

FPC faible... ou fort

- Évaluation des éléments de la compétence
- Non-utilisation de grilles d'évaluation descriptives.

COMPOSITION DE LA NOTE

Ce concept réfère aux différentes façons d'accorder des points ou des cotes en vue d'une évaluation sommative ou certificative. Avec ce concept, on tente d'identifier avec précision les objets sur lesquels on donne des notes et dans quelle proportion par rapport à la note finale.

Composantes de la note les moins utilisées

- Évaluation par les pairs
- Coévaluation enseignant-étudiant
- Le progrès, sans utiliser un prétest et un post-test
- Résultat d'équipe ou individuel dans un tournoi

Composantes de la note les plus utilisées

- Usage inapproprié de la langue écrite (pénalité)
- Trop d'absences entraînent un échec
- Utilisation des critères de performance pour noter l'épreuve finale
- Attribution de points à chaque élément de la compétence

Formation par objectifs (FPO) vs formation par compétences (FPC)

FPO fort

- Utilisation des critères de performance pour noter l'épreuve finale (problème d'intégration)

FPC faible... ou fort

- Utilisation des critères de performance pour noter l'épreuve finale (problème d'intégration)
- Non-utilisation de critères d'évaluation de la compétence (basés sur l'intégration des éléments de la compétence)

QUALITÉS DES ÉVALUATIONS

Le concept se rapporte aux qualités généralement exigées en évaluation, telles la rigueur, la validité, la clarté, la discrimination, la standardisation. Il vise à identifier les biais possibles en évaluation, tels l'influence de la force de la classe, le jugement des collègues, la propension ou l'incapacité à faire échouer un étudiant, le comportement d'un étudiant, la distribution des notes en relation avec la courbe normale.

Faiblesses les moins fréquentes

- Les résultats passés d'un étudiant influencent ses futures évaluations
- La force de la classe influence mes évaluations
- Mes évaluations visent à comparer les étudiants entre eux

Qualités les plus fréquentes

- Mes instruments évaluent les objets d'évaluation prévus
- Souci d'équité
- Consignes d'évaluation connues par l'étudiant avant l'évaluation
- Descriptions des évaluations dans le plan de cours

Formation par objectifs (FPO) vs formation par compétences (FPC)

Pour la dimension *Qualités des évaluations*, il a été impossible de créer des indices FPC et FPO

HIGHLIGHTS

Goal of this Study

The goal of this study was to describe or characterize the evaluation practices of CÉGEP Physical Education teachers in order to help develop knowledge, awareness and skills in the area of evaluation.

This section provides a brief definition of the concepts, the most and least frequent evaluation practices, as well as any significant tendencies in the objective-based and competency based learning practices.

EVALUATION PROCEDURES

The evaluation procedures deal primarily with the object of the evaluation (i.e.: knowledge, skills, competencies, elements of the competency, result or product, process, integration, others). These procedures refer also to the evaluation methods, the context, the limits and constraints of the evaluation, as well as the number and types of evaluations. They also focus on or are influenced by the Institutional Policy for the Evaluation of Student Achievement, as well as Departmental Regulations.

Least Frequent Evaluation Procedures

- Double Threshold
- Non standardized tests of physical fitness
- In game performance tests
- Evaluations requiring memorisation
- Three evaluations or less per semester

Most Frequent Evaluation Procedures

- Evaluations of the competency based upon the elements of the competency and the performance criteria
- Weighting of criteria according to their importance
- Information given to students prior to the evaluation regarding evaluation procedures
- Evaluations respecting departmental instructions or guidelines
- Plagiarism (detection and penalties)

Objective-Based Learning (OBL) vs Competency-Based Learning (CBL)

Strong OBL Tendency

- Evaluations weighted 12 % or less
- More than 4 evaluations per semester
- Less than 40% of the grades for the final evaluation

Weak CBL Tendency

- 3 evaluations or less per semester
- 40% or more of the grades for the final evaluation

MEASUREMENT TOOLS

Type of tests and exams used, types of evaluation tasks, methods of gathering information and the use of measurement tools

Least Used Measurement Tools

- Interviews
- Choice of Questions
- Oral Presentations

Most Used Measurement Tools

- Logbook
- Physical Fitness Tests
- Evaluation of the competency using integrative tasks
- Integrative tasks presented in the Course Outline
- Short answer written theory exams
- Integrative activity or test for entire course

Objective-Based Learning (OBL) vs Competency-Based Learning (CBL)

Strong OBL Tendency

- Written Theory Exams
- Short-answer questions

Weak CBL Tendency

- Le choix des instruments de mesure est bon (cahier de bord ou tâche d'intégration).
- Choice of measurement tools is good (logbook or integrative task)
- Cependant, on doute de la capacité des instruments de mesure à bien intégrer les éléments de compétence en un tout
- However, there is doubt as to whether the measurement tools truly integrate all the elements of the competency.

JUDGEMENT TOOLS

Through the use of measurement tools, judgement tools enable you to make a judgement (i.e.: to assign a grade or rating) as to the attainment of a goal or competency. These criteria are the performance criteria, the evaluation criteria, the elements of the competency and different types of grids or scales (rubrics).

Least used judgement tools

- I grade using a rating or scale.
- Criteria other than performance criteria
- Selection of only certain criteria of the competency

Most Used Judgement Tools

- Evaluation criteria selected by the department
- Elements of the competency used as criteria
- Use of all performance criteria of the competency

Objective-Based Learning (OBL) vs Competency-Based Learning (CBL)

Strong OBL Tendency

- Evaluation of all the performance criteria separately (integration problem)

Weak or Strong CBL Tendency

- Evaluation of the elements of the competency (Problem: How is this done?)
- Non-use of descriptive evaluation grids (rubrics)

GRADE MAKE-UP

This concept refers to the different ways of awarding marks or ratings for a summative evaluation. Efforts are made to precisely identify the objects or elements for which marks are awarded and in what proportion of the final grade.

Least used components of the grade make-up

- Peer evaluation
- Teacher-Student co-evaluation
- Improvement or progress without the use of a pre- and post-test
- Team or individual results in tournament

Most used grade make-up

- Inappropriate use of written language (deductions)
- Too many absences bring about a failure over-cut
- Use of performance criteria in grading the final task
- Awarding of marks for each element of the competency

Objective-Based Learning (OBL) vs Competency-Based Learning (CBL)

Strong OBL Tendency

- Utilisation des critères de performance pour noter l'épreuve finale (Problème d'intégration).
- The use of performance criteria in grading the final task (integration problem)

Weak or Strong CBL Tendency

- The use of performance criteria in grading the final task (integration problem)
- The non-use of evaluation criteria (based upon the integration of the elements of the competency)

-

QUALITY OF EVALUATIONS

This concept refers to the qualities generally required in evaluation, such as rigor, validity, clarity, discrimination and standardization. The aim is to identify possible evaluation biases, such as the influence of class strength or level, colleagues' judgement, the inclination or inability to fail a student, student behaviour, or the distribution of grades with respect to the normal curve.

Most Frequent Weaknesses

- Past student results influencing future evaluations
- The level or strength of the class influencing my evaluations
- My evaluations aim to compare students amongst themselves.

Most Frequent Qualities or Strengths

- My tools evaluate the expected objects of the evaluation
- Concerns regarding fairness or equity?
- Evaluation instructions known by students prior to the evaluation
- Description of the evaluation in the course outline

Objective-Based Learning (OBL) vs Competency-Based Learning (CBL)

Pour la dimension *Qualités des évaluations*, il a été impossible de créer des indices FPC et FPO.

It was impossible to detect strong or weak tendencies regarding OBL and CBL for this dimension of the study.

TABLE DES SUJETS

1. INTRODUCTION	2
2. DÉFINITION DU PROBLÈME	4
3. OBJECTIFS DE LA RECHERCHE	6
4. MÉTHODOLOGIE	7
4.1 Cadre de référence	7
4.2 Définition des concepts	8
4.3 Élaboration du questionnaire	10
5. RÉSULTATS	16
5.1 Modalités des évaluations	16
5.2 Instruments de mesure	28
5.3 Instrument de jugement	36
5.4 Composition de la note	40
5.5 Qualités des évaluations	46
6. CONCLUSION	52
ANNEXES	

1. INTRODUCTION

En mai 2010, le MELS a donné comme mandat à la coordonnatrice provinciale le mandat, pour l'année 2010-2011, de « ...faire un sondage sur les moyens utilisés par les départements pour constater la progression dans l'atteinte de l'un ou l'autre des devis ». Cette demande du ministère se voulait une façon de vérifier si des cégeps évaluaient le progrès, pratique qui, selon eux, ne respectait pas les devis ministériels. Afin de répondre à cette demande, un sondage a été réalisé à l'automne 2010 auprès des différents départements d'éducation physique de la province. Les résultats de ce sondage ont été présentés au comité provincial d'éducation physique au Collège Brébeuf, le 30 mai 2011 par Francine Lamarre et Luc Chiasson : *Sondage sur les difficultés de mise en œuvre des devis ministériels, version 2009*. À cette même occasion, François Lasnier présentait les modalités d'un autre sondage qui aurait comme but de préciser le sondage sur les difficultés de mise en œuvre en décrivant plus précisément les pratiques d'évaluation des enseignants. Comme ce mandat dépassait celui donné par le MELS, on a demandé à la Fédération des éducateurs et éducatrices enseignants du Québec (FÉÉPEQ) de parrainer ce projet. Pour cette organisation, l'évaluation des apprentissages est une partie importante et essentielle de la tâche d'enseignant et décrire les pratiques d'évaluation dans les collèges du Québec leur apparaissait un excellent moyen d'amorcer un processus d'amélioration des compétences des éducateurs physiques du niveau collégial en matière d'évaluation des apprentissages.

Pour le simple observateur, même si l'éducation physique est présente dans la formation générale obligatoire des cégeps depuis 1969, il apparaît que les modes d'évaluation en éducation physique ont très peu changé depuis plus de quarante ans. Ce qu'on peut observer, c'est que l'évaluation des apprentissages en éducation physique au niveau collégial montre très peu d'ouverture aux différents courants de pensée et semble surtout se développer à partir des expériences avec les étudiants, ainsi qu'avec les échanges entre pairs d'un même département ou de différents comités matières.

Le passage récent (1994) d'un programme d'éducation physique par objectifs vers un programme par compétences, aurait dû amener les enseignants à développer une nouvelle organisation de l'enseignement et de l'évaluation des apprentissages. Mais est-ce que ce fut le cas? Dans cette étude, nous allons vérifier dans quelle mesure les modes d'évaluation éducation physique collégiale ont été transformés par cette réforme.

2. DÉFINITION DU PROBLÈME

Dans les cégeps, la responsabilité de mesurer et d'évaluer les apprentissages des élèves en éducation physique relève des départements et des enseignants. Ces derniers sont généralement autonomes et ne sont pas assujettis à des règles précises. Les évaluations adoptées et les notes attribuées sont les seules qui prévalent dans la sanction des études et pour le calcul de la cote de rendement.

Pour beaucoup d'enseignants, la mesure et l'évaluation des apprentissages sont des sources de stress et de préoccupations. Il est très difficile de bien mesurer et de bien évaluer, et cette difficulté s'accroît encore lorsque vient le moment d'attribuer la note finale, principalement lorsqu'il ne s'agit plus seulement d'accompagner l'élève, mais de le juger et de certifier ses apprentissages. Si le jugement est si difficile, c'est principalement que les cibles d'apprentissage ne sont pas suffisamment claires, et qu'en premier lieu, la compréhension des compétences, des éléments et des critères de performance n'est pas suffisamment univoque. Ensuite, viennent les problèmes techniques relatifs aux types d'évaluation (formative ou sommative), à la construction et à l'utilisation des instruments de mesure, des grilles d'évaluation, des grilles de jugement et finalement de la pondération et de la notation des résultats.

Si autrefois des cours en mesure et évaluation obligatoires permettaient aux futurs enseignants de s'initier aux différentes tâches d'évaluation, il semble aujourd'hui que ces cours sont de moins en moins présents dans les programmes de formation initiale, ce qui place les nouveaux enseignants, dans une situation où évaluer relève plus de l'intuition que de la raison. Bien qu'à l'embauche, on requiert des nouveaux enseignants qu'ils aient une formation en éducation physique et en pédagogie, on n'exige pas d'eux qu'ils aient une formation adéquate en enseignement par compétences au niveau collégial. Cette formation pourrait les guider lors de l'évaluation des apprentissages de leurs étudiants. Le manque de spécialistes en mesure et évaluation dans les universités du Québec, ainsi que la fusion des départements de mesure et évaluation avec d'autres départements n'aide pas à améliorer la situation et est à bien des égards significative de l'intérêt qu'on porte à ce

champ de connaissance qui est pourtant indispensable pour démontrer la qualité de l'enseignement et des apprentissages.

Le savoir-faire collectif des éducateurs physiques en évaluation des apprentissages est très incertain puisque la formation en mesure et évaluation est inégale d'un professeur à l'autre et aussi à cause du caractère facultatif de la participation des enseignants à des activités de perfectionnement ou de formation continue.

Lorsqu'on a conscience de l'immense enjeu des pratiques d'évaluation sur toutes les dimensions du processus d'enseignement et des apprentissages, il devient important de clarifier comment contribuer à améliorer les pratiques des enseignants à l'égard de la mesure et de l'évaluation et ainsi protéger une liberté pédagogique tellement précieuse. À cet égard, il serait important d'identifier les pratiques des enseignants relatives à l'évaluation des apprentissages et de tenter de dégager des pistes de perfectionnement dans ce domaine.

Dans le but d'identifier des solutions aux problèmes liés à l'évaluation et à planifier des perfectionnements, un groupe d'enseignants appuyé par la FÉÉPEQ, a réalisé un sondage sur les pratiques évaluatives des enseignants et enseignantes en éducation physique du réseau collégial du Québec.

3. OBJECTIFS DE LA RECHERCHE

Objectif général :

L'objectif général de cette enquête est de décrire les pratiques d'évaluation des apprentissages des enseignants en éducation physique des collèges d'enseignement général et professionnel u Québec (cégeps) afin de les aider à s'améliorer. Cette description permettra d'évaluer le degré d'écart ou de conformité avec les principes d'évaluation en formation par compétences (FPC) et par objectifs (FPO), de façon à consolider ce qui est valable et à proposer des avenues de perfectionnement qui permettront d'améliorer les points faibles.

Objectifs spécifiques :

- Décrire les MODALITÉS D'ÉVALUATION utilisées dans les cours d'éducation physique au collégial.
- Décrire la nature et les applications des INSTRUMENTS DE MESURE.
- Décrire la nature et les applications des INSTRUMENTS DE JUGEMENT.
- Décrire quels sont les TYPES D'ÉVALUATIONS utilisés pour mesurer les critères de performance.
- Décrire la nature et les différentes COMPOSITIONS DE LA NOTE.
- Identifier les QUALITÉS DES ÉVALUATIONS.

4. MÉTHODOLOGIE

4.1 Cadre de référence

Le cadre de référence pour élaborer les instruments de mesure utilisés dans ce sondage est peu exhaustif. Il n'était aucunement dans nos intentions de produire une revue de la littérature en vue d'établir « l'état de la question » comme on peut lire dans la plupart des recherches. Nos points de référence reposent principalement sur l'expérience des auteurs en tant qu'enseignants et sur l'expertise en mesure et évaluation de certains des auteurs. Un des auteurs se spécialise depuis plus de vingt en élaboration d'instruments de mesure. C'est donc la mise en commun des connaissances et des expertises de chacun qui a permis de dresser un inventaire le plus exhaustif possible des pratiques évaluatives que nous avons observées en général au collégial, et de façon plus spécifique en éducation physique.

Les documents consultés sont peu nombreux étant donné le peu de publications spécifiques sur les inventaires des pratiques évaluatives au collégial. Nous nous sommes donc limités à deux ouvrages pour compléter notre liste de pratiques en évaluation. Ces documents sont :

Howe, R. et Ménard, L. (1993). *Croyances et pratiques en évaluation des apprentissages : étude des croyances et des pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages*. Rapport de recherche PAREA, Collège Montmorency.

Bélanger, D.-C., et Tremblay, K. (2012). *Portrait actualisé des croyances et des pratiques en évaluation des apprentissages au collégial*. Montréal : Collège Maisonneuve - PERFORMA.

La liste des pratiques évaluatives que nous avons établie pour ce sondage ne se limite pas aux bonnes pratiques seulement ou aux pratiques évaluatives compatibles avec les programmes définis par compétences. Bien que nous ayons pris en compte les pratiques associées à l'évaluation des compétences, nous avons aussi énuméré les

pratiques qu'on observe dans les programmes définis par objectifs. Cette façon de procéder s'imposait puisqu'on observe encore de nombreuses pratiques évaluatives associées à la formation par objectifs.

4.2 Définition des concepts

Les regroupements des pratiques, pour constituer des concepts en évaluation, sont les suivants :

- MODALITÉS D'ÉVALUATION
- INSTRUMENTS DE MESURE
- INSTRUMENTS DE JUGEMENT
- COMPOSITION DE LA NOTE
- QUALITÉS DES ÉVALUATIONS

MODALITÉS D'ÉVALUATION

Les modalités d'évaluation, appliquées à l'évaluation sommative ou certificative, concernent principalement les objets d'évaluation (connaissances, habiletés, compétences, éléments de compétence, produit, processus, intégration, autres). Les modalités d'évaluation réfèrent aussi aux méthodes d'évaluation, au contexte, aux limites et contraintes, au nombre et aux types d'évaluation. L'évaluation des modalités porte aussi sur la *Politique institutionnelle d'évaluation des apprentissages* et les règles départementales.

INSTRUMENTS DE MESURE

Étant donné la généralisation de l'implantation des programmes définis par compétences, nous avons jugé important de bien distinguer les instruments de mesure et les instruments de jugement. Dans ce questionnaire d'auto-évaluation, le concept relatif aux instruments de mesure réfère aux types de tests et d'examens utilisés, aux types de tâches d'évaluation (travaux longs, études de cas, séries de questions, types de question, tâches d'intégration, examen-synthèse, autres). Le concept touche aussi aux modes de cueillettes d'information (fiches, rapports, portfolio, carnet de bord, autres), ainsi qu'aux

modes d'application des instruments (individuel, en équipe, production écrite ou verbale, démonstration, production d'une oeuvre, autres).

INSTRUMENTS DE JUGEMENT

Les instruments de jugement diffèrent des instruments de mesure, car ils ne servent pas à recueillir des informations, mais à porter un jugement à partir des instruments de mesure (*Exemple : je recueille des informations à partir d'une tâche ou d'un portfolio, mais je note ou je cote à partir d'une grille d'évaluation ou une liste de critères d'évaluation. Même pour un examen traditionnel, composé d'une série de questions déjà pondérées, je peux construire un instrument de jugement pour noter chaque réponse en fonction d'un résultat attendu.*) Dans ce questionnaire, les instruments de jugement concernent les critères de performance, les critères d'évaluation, les éléments de la compétence, les différents types de grilles, dont les grilles descriptives, avec des échelons correspondant à différents niveaux de développement de la compétence évaluée.

COMPOSITION DE LA NOTE

Le concept *Composition de la note* réfère aux différentes façons d'accorder des points ou des cotes en vue d'une évaluation sommative ou certificative. Avec ce concept, on tente d'identifier avec précision les objets sur lesquels on donne des notes et dans quelle proportion par rapport à la note finale. Le concept couvre aussi le ou les évaluateurs potentiels (l'enseignant, l'étudiant, les pairs, l'équipe de travail). La composition de la note couvre aussi certains aspects particuliers, tels l'évaluation du progrès, des performances, les absences, les attitudes, les comportements et la qualité de la langue.

QUALITÉS DES ÉVALUATIONS

Ce concept concerne la concordance entre les objets d'évaluation, le plan de cours et la nature des évaluations. Le questionnaire considère autant les instruments de mesure que les instruments de jugement. Le concept se rapporte aux qualités généralement exigées en évaluation, telles la rigueur, la validité, la clarté, la discrimination, la standardisation. Le questionnaire vise à identifier les biais possibles en évaluation, tels l'influence de la force de la classe, le jugement des collègues, la propension ou

l'incapacité à faire échouer un étudiant, le comportement d'un étudiant, la distribution des notes en relation avec la courbe normale.

4.3 Élaboration du questionnaire

Les travaux du comité se sont échelonnés sur une période de près de deux ans. Voici les principales étapes avec une brève explication de chacune d'elle.

Hiver 2011 : Construction du questionnaire et traduction.

Automne 2011 : Validation du questionnaire et traduction.

Novembre 2011 : Envoi du questionnaire.

Janvier 2012 : Analyse des résultats.

Mai 2012 : Présentation des résultats au comité provincial à La Pocatière.

Février 2013 : Rapport d'analyse du sondage.

CONSTRUCTION DU QUESTIONNAIRE ET TRADUCTION

L'identification des items du questionnaire s'est faite à partir des concepts modalités d'évaluation, instruments de mesure, instruments de jugement, composition de la note et qualités des évaluations et de chacun des sous-concepts des différents concepts. Voir le tableau 4.1 pour un exemple sur la façon d'identifier les items à partir d'un concept et ses sous concepts

La traduction du questionnaire a été assurée par M. Bob Delaney, enseignant retraité du Cégep Champlain – Saint-Lawrence. Il a assuré pendant plusieurs années la traduction des principaux documents du comité provincial d'éducation physique et du MELS concernant l'éducation physique.

Tableau 4.1 Exemple d'identification d'items du questionnaire, selon un concept et un sous-concept.

Concept : Composition de la note	
Sous-concepts	Items
Notation de chaque critère de performance	J'accorde des points à chaque critère de performance de la compétence.
Notation de l'épreuve finale	J'utilise les critères de performance de la compétence pour noter l'épreuve finale.
	J'utilise les éléments de la compétence pour noter l'épreuve finale.
	J'utilise mes propres critères d'évaluation pour noter l'épreuve finale.

ÉLABORATION DES INSTRUMENTS DE MESURE

Pour l'élaboration de l'instrument de mesure de chacun des concepts pris en compte pour classer les pratiques évaluatives, les membres du comité de rédaction, composé par trois des auteurs, ont d'abord sélectionné les regroupements sous lesquels il semblait le plus pertinent de classer les pratiques. Ces regroupements sont : modalités d'évaluation, instruments de mesure, instruments de jugement, composition de la note et qualités des évaluations.

Puis, pour chacun de ces regroupements, nous avons analysé les concepts afin d'en déduire les sous-concepts. Une fois cette opération effectuée, chaque membre du comité de rédaction écrivait deux items (énoncé ou question) représentant des pratiques évaluatives associées à chaque sous-concept. À partir de cette liste d'items exhaustive, nous avons retenu les meilleurs items tout en recherchant une bonne validité de concepts, c'est-à-dire un nombre suffisant d'items pour couvrir tous les aspects de chaque sous-concept. Pour réaliser la sélection des items, nous cotions chaque item avec les cotes variant de 1 à 3 (3 étant bon et 1 étant faible). Pour coter la qualité des items, nous tenions compte de la clarté du contenu, de la qualité de sa rédaction et de sa pertinence en fonction de la nature du concept mesuré.

VALIDATION PAR UN COMITE DE D'EXPERTS

Une fois la sélection des items terminée, nous avons procédé à la validation de chaque instrument en utilisant un comité d'experts. Trois des quatre membres de ce comité étaient des enseignants du collégial qui possédaient une longue expérience en enseignement et qui avaient manifesté un intérêt marqué pour la problématique de l'évaluation des compétences en éducation physique. Chaque membre du comité d'experts cotait chaque item selon deux critères : la pertinence et la formulation. Chaque critère était évalué avec une échelle de mesure s'échelonnant de 1 à 6 ; 6 étant la pertinence et la qualité de la formulation la meilleure. Le tableau 4.2 illustre le modèle utilisé par le comité d'expert. La validation par le comité d'expert a conduit à l'élimination de quelques items qui n'obtenaient pas un score suffisamment élevé sur un des deux critères. De plus, l'évaluation des membres du comité d'expert a conduit à la correction de quelques items.

Tableau 4.2 Exemple du document permettant aux experts de valider les items du questionnaire.

Exemple : Composition de la note			
Sous-concept	Items	Pertinence	Formulation
Notation de chaque critère de performance	J'accorde des points à chaque critère de performance de la compétence.	1 2 3 4 5 6	1 2 3 4 5 6
Notation de l'épreuve finale	J'utilise les critères de performance de la compétence pour noter l'épreuve finale.	1 2 3 4 5 6	1 2 3 4 5 6
	J'utilise les éléments de la compétence pour noter l'épreuve finale.	1 2 3 4 5 6	1 2 3 4 5 6
	J'utilise mes propres critères d'évaluation pour noter l'épreuve finale.	1 2 3 4 5 6	1 2 3 4 5 6

VALIDATION AUPRÈS D'UN ÉCHANTILLON RESTREINT

Une fois la version des instruments de mesure établie, nous avons procédé à une expérimentation auprès d'un échantillon restreint. Deux membres du comité de rédaction ont rencontré individuellement six enseignants qui venaient juste de prendre leur retraite (nous avons utilisé des retraités pour ne pas perdre de sujets dans le sondage). Nous avons demandé à chaque enseignant ou enseignante, sélectionné pour cette étape de pré expérimentation, de répondre à tous les items et de signaler tout problème de compréhension rencontré à la lecture des items et aussi de nous indiquer s'il y avait un problème à répondre à certains items, ou même si cette personne avait des suggestions quant à la formulation d'un item qu'elle avait trouvé ambiguë. Cette opération a contribué à corriger quelques items et à nous confirmer que le questionnaire était suffisamment compréhensible pour être appliqué sur un grand échantillon.

MISE EN LIGNE

Le questionnaire a été mis en ligne et envoyé à l'adresse courriel de chaque enseignant en éducation physique des cégeps publics et privés du Québec. Pour ce faire nous avons utilisé le logiciel Netsondage élaboré par le Centre collégial de développement de matériel didactique CCDMD. Il a été déposé sur le site de Prof web (hébergement et support) grâce à l'aide de M. Philippe Aubé (coordonnateur profweb) et à l'appui technique de M. Jonathan Marc Lapointe (technicien profweb).

NOMBRE DE RÉPONDANTS

Le tableau 4.3 présente le nombre de répondants pour chaque ensemble du programme en éducation physique. Un total de 131 répondants sur un total possible d'environ 450, soit environ 30 %, ont répondu au questionnaire. Il aurait été souhaitable de recevoir plus de questionnaires, mais nous en avons suffisamment pour établir des observations intéressantes et tirer des conclusions utiles.

Tableau 4.3 Nombres de répondants par ensemble.

Ensemble	N ^{bre}	%
Ensemble 1	58	44,3
Ensemble 2	45	34,4
Ensemble 3	28	21,4
Total	131	100,0

Nombre possible de répondants : environ 450

CUEILLETTE DES DONNÉES

La cueillette des données pour identifier les pratiques évaluatives a été réalisée au moyen d'un questionnaire présentant une vaste énumération des pratiques possibles en évaluation pour les éducateurs physiques. Chacun des concepts devait être évalué sous deux aspects : la fréquence des pratiques et la proximité avec la formation par compétences (FPC) et la formation par objectifs (FPO).

L'inventaire des pratiques en évaluation constitue une liste des pratiques qu'il est possible d'observer dans les cégeps du Québec. Certaines pratiques peuvent être fréquentes dans une institution et inexistantes dans une autre. Cette section du questionnaire vise d'abord à identifier les pratiques qui sont les plus et les moins populaires.

ANALYSES ET TRAITEMENTS DES DONNÉES

Dans un premier temps, nous avons procédé à des analyses exploratoires des données. Les analyses exploratoires du présent projet poursuivent principalement les objectifs suivants : vérifier le codage des données, mieux comprendre la nature de l'échantillon, mieux comprendre les données à analyser, mieux comprendre les concepts étudiés. Cette forme d'analyse se fait selon certaines stratégies et certaines techniques statistiques reconnues en théorie de la mesure (Tukey, 1977; Bertrand, 1986; Lasnier, 1989). Les données ont été traitées avec le logiciel de statistique SPSS.

Dans un deuxième temps, nous avons réalisé une analyse des résultats item par item. Pour chaque item, nous avons calculé la distribution des fréquences, la moyenne, et l'écart type afin de valider les données et de réaliser une première analyse de chaque item. Cette opération nous a permis d'identifier les pratiques évaluatives les plus fréquentes et les moins fréquentes.

La seconde partie de l'analyse des données consistait à étudier chaque concept relatif à l'évaluation (*modalités d'évaluation, instruments de mesure, instruments de jugement, composition de la note et qualités des évaluations*) en fonction de sa proximité avec la formation par compétences ou la formation par objectifs. Nous avons donc tenté de construire une mesure unidimensionnelle de chacun des concepts. Nous n'avons pas réussi à créer des mesures suffisamment unidimensionnelles pour établir des concepts cohérents. Nous avons donc libellé ces mesures des « indices » plutôt que des concepts. Il n'est pas vraiment surprenant que les regroupements planifiés ou concepts n'arrivent pas à des mesures unidimensionnelles étant donné la grande disparité dans la liste des pratiques évaluatives qui constitue l'inventaire présenté dans le questionnaire. En effet, on y retrouve des bonnes et des mauvaises pratiques, des pratiques associées à la FPO et d'autres associées à la FPC. Toutefois, en faisant une sélection d'items parmi les items de chaque concept analysé, nous avons été capables de créer des indices nous permettant d'analyser les pratiques évaluatives en fonction de leur proximité avec la FPO et la FPC.

5. RÉSULTATS

Chacun des concepts a fait l'objet d'une analyse d'items selon la théorie classique des tests, afin d'en analyser la validité et sa fidélité en terme de la cohérence interne. Lorsqu'il sera impossible de créer une mesure unidimensionnelle du concept, nous composerons des indices de mesure afin de pouvoir donner un aperçu de la nature de ces concepts, et ce, même si ces mesures sont moins valides qu'une mesure unidimensionnelle.

5.1 Modalités des évaluations

Les modalités d'évaluation, appliquées à l'évaluation sommative ou certificative, concernent principalement les objets d'évaluation (connaissances, habiletés, compétence, éléments de compétence, produit, processus, intégration, autres). Les modalités d'évaluation réfèrent aussi aux méthodes d'évaluation, au contexte, aux limites et contraintes, au nombre et aux types d'évaluation. L'évaluation des modalités porte aussi sur la *Politique institutionnelle d'évaluation des apprentissages* et les règles départementales.

Le tableau 5.1 présente la moyenne et l'écart type de chaque item du questionnaire pour la section 2 : *Modalités d'évaluation*. Les items sont présentés dans un ordre croissant selon les moyennes. Le tableau permet de visualiser les pratiques les plus fréquentes et les moins fréquentes. Il ne faut pas oublier que la moyenne doit s'interpréter à la lumière de l'écart type.

Notons que le concept *Modalités d'évaluation*, n'est pas une mesure unidimensionnelle (ensemble d'items reliés entre eux), mais plutôt un indice, car il a été impossible de créer un ensemble d'items interreliés et pour lesquels nous pourrions calculer un coefficient de cohérence interne, comme par exemple, le coefficient alpha. Cela se produit fréquemment lorsque les items réfèrent à un inventaire de pratiques pédagogiques, car certaines pratiques sont en opposition avec d'autres, ou encore, certaines pratiques peuvent être bonnes, mais ne pas être utilisées dans certains contextes.

Tableau 5.1 Moyenne et écart type des items de la section *Modalités d'évaluation*

Item	Moyenne	Écart type
Q2.32	1,46	1,15
Q2.33	1,87	1,62
Q2.31	2,18	1,81
Q2.9	2,24	1,72
Q2.26	2,54	2,08
Q2.14	2,76	1,40
Q2.28	2,82	1,58
Q2.7	2,95	2,07
Q2.29	3,15	1,73
Q2.15	3,24	1,99
Q2.35	3,24	2,20
Q2.6	3,42	2,07
Q2.23	3,83	1,67
Q2.3	3,95	2,02
Q2.4	3,95	1,95
Q2.8	4,16	2,17
Q2.5	4,29	1,62
Q2.11	4,30	1,73
Q2.24	4,30	1,62
Q2.2	4,44	1,72
Q2.1	4,54	1,58
Q2.34	4,59	1,63
Q2.18	4,68	1,31
Q2.17	4,70	1,34
Q2.16	4,79	1,41
Q2.27	4,98	1,76
Q2.13	5,05	,93
Q2.10	5,16	1,22
Q2.12	5,18	1,06
Q2.30	5,20	1,04
Q2.25	5,37	1,21
Q2.21	5,56	,92
Q2.22	5,69	,63
Q2.20	5,70	,57
Q2.19	5,78	,51
Moyenne	4,05	

Note. Les numéros des items en gras indiquent une différence significative entre les ensembles (Q2.1, Q2.4, Q2.6, Q2.7, Q2.8). C'est essentiellement l'ensemble 2 que est différent des 2 autres ensembles. Cela semble cohérent d'après la nature des compétences.

5.1.1 Les modalités d'évaluation les moins fréquentes

Items relatifs à l'épreuve finale et la règle du double seuil

- Q2.31 Pour mon évaluation, j'applique la règle du **double seuil** suivante : 60 % sur l'ensemble des évaluations (incluant l'épreuve finale) et en plus avoir au moins 60 % à l'épreuve finale.
- Q2.32 Pour mon évaluation, j'applique la règle du **double seuil** suivante : avoir déjà obtenu 60 % pour accéder à l'évaluation finale et obtenir au moins 60 % à l'évaluation finale.
- Q2.33 Pour mon évaluation, si un étudiant n'obtient pas 60 % pour l'ensemble de ses évaluations, mais qu'il **obtient 60 % et plus à l'épreuve finale**, il réussit automatiquement le cours.

Globalement, on peut affirmer que, chez nos répondants, une règle du double seuil lors des évaluations est rarement appliquée. Selon ces résultats, il semble qu'on exige que *très rarement* ou *jamais* que l'étudiant ait :

- Avoir obtenu 60 % pour l'ensemble de la session pour se présenter à l'examen final.
- Avoir obtenu 60 % à l'ensemble de ses évaluations en plus d'avoir réussi son épreuve finale.
- Avoir seulement réussi son examen final pour réussir le cours.

En considérant que les réponses à la question Q2.26 (*Je fais 3 évaluations et moins par session*) et Q2.35 (*Mon épreuve synthèse de cours (épreuve finale) compte pour au moins 40 % des points de la session*), nous indiquent que les enseignants font beaucoup d'évaluations pendant la session (c.-à-d. : plus que 4, voir question Q2.27) et que l'épreuve finale compte généralement pour moins de 40 points. Cet ensemble d'informations nous permet d'affirmer que la majorité des répondants évaluent selon une FPO et non pas selon une FPC qui privilégie de réduire le nombre d'évaluations sommatives en cours de session (3 et moins), qui préconise aussi des évaluations formatives et une épreuve synthèse de cours comportant plus de points et d'importance (c.-à-d. : 40 % et plus des points de la session).

Items relatifs à l'utilisation de tests de condition physique.

- Q2.9 J'évalue la condition physique sans tests standardisés (tests sans normes de références).
- Q2.8 J'évalue la condition physique avec des tests standardisés (ex. : physitest canadien, redressements assis).

Selon nos répondants (Q2.9), il est très rare qu'on évalue le niveau de condition physique avec des tests sans normes de référence (ou avec des tests maison). On semble plutôt opter pour des tests standardisés. Cependant, un écart type de 2,07, nous indique que cette utilisation de tests standardisés peut varier beaucoup autour de la moyenne (*sur une échelle de 1 à 6*). La compréhension de la notion de *tests standardisés* n'est certainement pas étrangère à cette variabilité.

Items relatifs à l'utilisation des tests de performances physiques

- Q2.7 J'évalue les performances physiques en situation de jeu (ex. : application des stratégies collectives en jeu).
- Q2.6 J'évalue les performances physiques (ex. : nombre de lancers réussis, technique de la manchette au volleyball).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q2.6 (1-2 et 2-3)	2,4	1,9	5	1,4	2,9	1,9
Q2.7 (1-2 et 2-3)	2	1,6	4,6	1,8	2,4	1,8

M = moyenne; *s* = écart type

Note : Q2.6 (1-2 et 2-3) indique qu'il y a une différence statistiquement significative entre la moyenne obtenue pour l'ensemble 1 et l'ensemble 2 et entre la moyenne obtenue pour l'ensemble 2 et l'ensemble 3. Il n'y a pas cependant de différence entre la moyenne de l'ensemble 1 et de l'ensemble 3.

L'utilisation des tests de performance physique (en situation de jeu ou non) varie selon les ensembles. C'est dans l'ensemble 2 qu'on utilise le plus souvent ces tests. C'est normal que la moyenne soit élevée puisque le sens de la compétence de cet ensemble est

d'améliorer l'efficacité lors de la pratique d'une activité physique. Par ailleurs, il est étonnant de constater qu'on évalue les performances physiques, même si ce n'est pas beaucoup, dans l'ensemble 1 et dans l'ensemble 3 et qu'on le fait plus souvent dans l'ensemble 1 que dans l'ensemble 3. Ici encore, ce résultat est sans doute attribuable à des compréhensions différentes de la notion de « tests de performance physique ».

Items relatifs à l'évaluation dite « théorique »

- Q2.14 Je fais des évaluations qui font appel au par cœur.

Items	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q2.14 (2-3)	2,8	1,4	3	1,4	2,2	1,3

Il y a très peu de répondants qui font des évaluations faisant appel au par cœur. Si cette pratique est semblable pour l'ensemble 1 et 2, elle est différente pour l'ensemble 3 où elle est encore moins courante. Ce qui est normal étant donné la nature de l'ensemble 3.

Est-ce qu'une évaluation qui ne fait pas appel au par cœur nous indique qu'il n'y a pas d'exams théoriques ? Les réponses aux questions 2.1 et 2.5 (voir plus bas) nous indiquent qu'on semble faire une nette distinction entre *faire appel au par cœur* et évaluer les connaissances disciplinaires et les habiletés intellectuelles. Comme pour la question 2.4, c'est dans l'ensemble 3 qu'on évalue moins les connaissances disciplinaires. Pour les habiletés intellectuelles, il n'y a pas de différence statistiquement significative entre les moyennes des trois ensembles (dans l'ordre : 4,5, 4,4 et 3,7). Par ailleurs, il serait intéressant de vérifier quels sont les moyens utilisés pour évaluer les connaissances disciplinaires (Q2.1), les habiletés intellectuelles (Q2.5) ou le par cœur.

Q2.1 J'évalue les connaissances disciplinaires (savoirs, connaissances déclaratives) (moyenne 4,54 et écart-type de 1,62).

Q2.5 J'évalue les habiletés intellectuelles (moyenne 4,29 et écart-type de 1,58).

Items relatifs à la répartition de la valeur des évaluations pendant la session.

- Q2.26 Je fais 3 évaluations et moins par session. (moyenne 2,54 et écart-type de 2,084).
- Q2.28 Je répartis également les points de mon évaluation sur l'ensemble de la session (ex. : 50 % des points avant la mi-session et 50 % après la mi-session) (moyenne 2,82 et écart-type de 1,587).
- *Q2.27 Je fais 4 évaluations et plus par session (moyenne 4,98 et écart-type de 1,767).*

Très peu de répondants font 3 évaluations et moins par session. La réponse à la question Q2.27 nous précise qu'ils en font 4 et plus. Comme en formation par compétences (FPC) on suggère de réduire le plus possible le nombre d'évaluations sommatives et de surtout faire des évaluations formatives, ces deux réponses nous suggèrent qu'on fait plutôt des évaluations traditionnelles ou par objectifs. Toutefois, la moyenne de la question Q2.8 (2,82) nous indique que la majorité des points doivent se donner après la mi session, ce qui rejoint en partie la FPC qui suggère de laisser le plus de temps possible à l'étudiant pour développer sa compétence avant de l'évaluer sommairement.

5.1.2 Les modalités d'évaluation les plus fréquentes

Pour les items les plus forts, on remarque qu'il n'y a pas de différence statistiquement significative entre les moyennes des trois ensembles ou cours. Ainsi, nos remarques ne tiendront pas compte de différences possibles entre les ensembles 1, 2 ou 3.

Items relatifs à l'évaluation de la compétence

- Q2.13 Mes évaluations permettent de juger du degré de développement de la compétence (moyenne 5,05 et écart-type de 0,93).
- Q2.10 J'évalue les éléments de la compétence en utilisant les critères de performance tels qu'écris dans le devis ministériel (moyenne 5,16 et écart-type de 1,22).
- Q2.12 J'évalue tous les éléments de la compétence (moyenne 5,18 et écart-type de 1,06).
- *Q2.11 Je réfère aux éléments de la compétence pour rédiger mes propres critères d'évaluation (différents des critères de performance) (moyenne 4,30 et écart-type de 1,73).*

Globalement, les enseignants ont la conviction d'évaluer le degré de développement de la compétence. Ils évaluent les éléments de la compétence en utilisant les critères de performance proposés dans les devis. Ils affirment également évaluer tous les éléments de la compétence. Il faut signaler en se rapportant à la question Q2.11 qui est moins populaire, que certains enseignants évaluent les éléments de compétence avec leurs propres critères de performance qui peuvent être différents des critères de performance du devis. Est-ce qu'en évaluant tous les critères de performance de la compétence, on évalue par incidence tous les éléments de la compétence et la compétence considérée comme un tout. Ce n'est pas sûr. Nous tenterons de répondre à cette question plus loin dans cette section (voir Analyse FPO versus FPC) et également dans la section sur les instruments de jugement.

Items relatifs à la pondération des critères d'évaluation.

- Q2.30 Je pondère mes critères d'évaluation selon leur importance (moyenne 5,20 et écart-type de 1,04).

Une très grande majorité de nos répondants pondèrent leurs critères d'évaluation selon leur importance. Ainsi, on n'accorde pas la même importance à tous les critères, ce

qui semble une procédure normale puisqu'à la lecture du devis, les critères n'ont pas la même importance. Cette importance se traduisant par le temps qu'on accorde aux activités d'apprentissage.

Items relatifs aux informations données aux étudiants par rapport à l'évaluation

- Q2.20 J'informe mes étudiants, à l'avance, des critères d'évaluation (moyenne 5,70 et écart-type de 0,57)
- Q2.21 J'informe mes étudiants, à l'avance, de la pondération des critères d'évaluation (moyenne 5,56 et écart-type de 0,92).
- Q2. 19 J'informe mes étudiants, à l'avance, sur quoi portera l'évaluation (moyenne 5,78 et écart-type de 0,51).

Parmi les items relatifs aux modalités d'évaluation, ce sont ceux qui se rapportent à l'information donnée aux étudiants sur les évaluations qui ont obtenu le meilleur résultat. Ainsi, lors des évaluations, la majorité des répondants informent leurs étudiants des critères d'évaluation et de leur pondération, ainsi que sur le contenu de cette l'évaluation (le quoi).

Items relatifs à la directive du département lors des évaluations

- Q2.22 Je me conforme aux directives du département lors de mes évaluations.

La majorité des enseignants affirment se conformer aux directives du département lors des évaluations. Notre questionnaire ne permet pas de préciser le genre et le niveau de directives que le département donne à ses enseignants.

Concernant le plagiat dans le cours d'éducation physique

On observe qu'il y a de plus en plus de plagiat dans les cégeps et probablement aussi en éducation physique. Sans vraiment savoir les types de plagiat qui sont utilisés

en éducation physique (*ce n'était pas le but de cette étude*) il est intéressant de savoir que la majorité de nos répondants sanctionnent tout plagiat.

- Q2.25 Je sanctionne tout plagiat (moyenne 5,37 et écart-type de 1,21).

Cependant, bien qu'on affirme sanctionner tout plagiat, beaucoup de répondants ne semblent pas nécessairement avoir mis en place de moyens systématiques pour les prévenir ou les détecter.

- Q2.24 Je mets en place des moyens pour prévenir le plagiat (moyenne 4,30 et écart-type de 1,62).
- Q2.23 J'utilise des moyens de détection des cas de plagiat (moyenne 3,83 et écart-type de 1,67).

5.1.3 Modalités d'évaluation : Analyse FPO versus FPC

Pour le concept *Modalités des évaluations*, nous avons regroupé deux ensembles d'items qui constituent des indices de FPC et de FPO, et nous avons calculé la moyenne et l'écart type de chacun des deux regroupements d'items. Ceux-ci ne sont pas unidimensionnels (*items fortement interreliés*), c'est pourquoi nous les nommons des indices. Examinons s'il y a une différence entre ces deux regroupements. Étant donné qu'il n'y a pas de différence significative entre les trois ensembles sur les items qui composent ces indices, on ne sépare pas les ensembles pour cette analyse. Nous vous présentons les items qui ont constitué le regroupement FPO et le regroupement FPC. Le tableau 5.2 présente la moyenne et l'écart type pour les indices FPC et FPO.

Indices FPC

- Q2.11 Je réfère aux éléments de la compétence pour rédiger mes propres critères d'évaluation (différents des critères de performance).
- Q2.12 J'évalue tous les éléments de la compétence.

- Q2.15 Lors de mes évaluations, les étudiants ont accès à leurs notes de cours et aux autres documents utilisés dans le cours.
- Q2.26 Je fais 3 évaluations et moins par session.
- Q2.35 Mon épreuve épreuve-synthèse de cours (épreuve finale) compte pour au moins 40 % des points de la session.

Indices FPO

- Q2.10 J'évalue les éléments de la compétence en utilisant les critères de performance tels qu'écrits dans le devis ministériel.
- Q2.14 Je fais des évaluations qui font appel au par cœur.
- Q2.27 Je fais 4 évaluations et plus par session.
- Q2.28 Je répartie également les points de mon évaluation sur l'ensemble de la session (ex. : 50 % des points avant la mi-session et 50 % après la mi-session).
- Q2.34 Je fais certaines évaluations qui comptent pour moins de 12 % des points de la session.

Tableau 5.2 Moyenne et écart type des indices FPC et FPO (Modalités d'évaluation).

Variable	<i>n</i>	<i>M</i>	<i>s</i>
C1MoyFPC	131	3,7	0,86
C1MoyFPO	131	4,0	0,81

n = nombre de répondants; *M* = moyenne; *s* = écart type

On observe une différence significative entre la moyenne des items FPC et celle des items FPO ($F_{(1, 130)} = 9,53$; $p = 0,002$). La moyenne FPO est significativement plus élevée que celle FPC. Cela signifie que les enseignants appliquent davantage la FPO que la FPC pour les **modalités de l'évaluation**.

Voici un autre exemple à noter : Q2.26 (3 évaluations et -) ($M = 2,54$) versus Q2.27 (4 évaluations et +) ($M = 4,98$). Un tel exemple montre bien que les enseignants sont davantage orientés vers la FPO que la FPC quant aux modalités d'évaluation. Autre exemple : Q2.34 (12 % et -) ($M = 4,59$) versus Q2.35 (40 % et +)

($M = 3,24$). Un tel exemple montre encore que les enseignants sont davantage orientés vers la FPO que la FPC. Toutefois, dans l'ensemble 3, les enseignants sont moins dans une approche d'évaluation centrée sur la FPO.

Nous vous présentons un autre exemple à analyser : Q2.14 (par cœur) (Moyenne = 2,74) versus Q2.15 (droit aux notes de cours) (Moyenne = 3,24). Ici, les enseignants montrent davantage une bonne pratique. Toutefois, il est très étonnant de constater que c'est dans l'ensemble 2 qu'on fait le plus d'évaluation qui font appel au par cœur.

L'exemple suivant nous interroge aussi : Q2.34 avec une moyenne de 4,59 ($s = 1,63$), (évaluation pour moins de 12%). C'est cohérent avec Q2.27 (4 évaluations et plus : moyenne = 4,98 ; $s = 1,76$). Toutefois, une moyenne de 4,59 pour Q2.34 (évaluation pour moins de 12%). demeure nettement une évaluation centrée sur la FPO, car elle est trop élevée. Elle devrait tendre vers 0 dans une évaluation centrée sur la FPC.

Le tableau 5.3 résume les principales caractéristiques d'une évaluation centrée sur la FPO comparativement à une évaluation centrée sur la FPC. Le tableau 5.3 dichotomise la FPO et la FPC pour en faciliter l'illustration. Cependant, dans la réalité, les observations qui y sont présentées sont davantage associées à l'un ou l'autre des deux types de formation. Le fait qu'une observation soit présentée dans une section n'exclut donc pas la possibilité de l'observer dans l'autre type de formation, mais elle se manifeste à une fréquence moins élevée.

Tableau 5.3 Caractéristiques de l'évaluation principalement associées à l'évaluation en FPO et en FPC.

Évaluation en FPO	Évaluation en FPC
Évaluation relativement facile	Évaluation exigeante
Mesure plutôt objective	Mesure plutôt subjective, mais davantage rigoureuse
Quelquefois, scissure entre apprentissage et évaluation	Recherche l'intégration entre enseignement, apprentissage et évaluation.
Sollicite peu l'évaluation formative	Sollicite beaucoup l'évaluation formative
Évaluation par questions et quelquefois par développements	Évaluation par tâche d'intégration qui sollicite tous les éléments de la compétence
Sélection du contenu par thèmes ou objectifs (cherche à couvrir l'ensemble du contenu disciplinaire)	Sélection du contenu en fonction de la tâche et de la situation reliée à la compétence
Évalue une seule fois un objet d'évaluation	Idéalement, mesure plusieurs fois un objet d'évaluation (itération)
Évaluation souvent normative : comparaison entre les étudiants	Évaluation critériée : comparaison des résultats avec des critères basés sur des résultats attendus.
Accorde beaucoup d'importance au produit (performance)	Accorde l'importance au processus (la démarche) et au produit
Critères d'évaluation absents ou peu définis	L'évaluation devrait se faire avec les critères d'évaluation de la compétence (à définir en département) et non avec les critères de performance
Souvent centrée sur le par coeur	On donne accès aux informations
Davantage quantitative	Davantage qualitative.
Questions indépendantes les unes des autres	Centrée sur l'intégration des éléments de compétence
Correction systématique	Jugement porté au moyen d'une grille descriptive avec des échelons, décrivant différents niveaux de maîtrise de la compétence
Informe sur la maîtrise des objectifs	Informe sur le degré de maîtrise de la compétence
Généralement, évaluation au moyen d'exams	Évaluation avec des tâches d'intégration
Généralement, notation question par question	Notation à partir de la grille descriptive, demande de porter un jugement
Évaluation sommative	Évaluation certificative

5.2 Instruments de mesure

Étant donné la généralisation de l'implantation des programmes définis par compétences, nous avons jugé important de bien distinguer les instruments de mesure et les instruments de jugement. Dans ce questionnaire d'auto-évaluation, le concept relatif aux instruments de mesure réfère aux types de tests et d'examens utilisés, aux types de tâches d'évaluation (travaux longs, études de cas, séries de questions, types de question, tâches d'intégration, examen-synthèse, autres). Le concept touche aussi aux modes de cueillettes d'information (fiches, rapports, portfolio, carnet de bord, autres), ainsi qu'aux modes d'application des instruments (individuel, en équipe, production écrite ou verbale, démonstration, production d'une œuvre, autres).

Les tableaux 5.4 et 5.5 présentent la moyenne et l'écart type pour les items de la section *Instruments de mesure*.

Tableau 5.4 Moyenne et écart type pour les items de la section *Instruments de mesure* (3 ensembles regroupés)

Item	Moyenne	Écart type
Q3.12*	1,51	1,14
Q3.18	1,60	1,28
Q3.10	1,82	1,43
Q3.8	2,44	1,78
Q3.19	2,56	1,81
Q3.9	2,66	1,88
Q3.17	2,88	1,83
Q3.3	2,89	2,01
Q3.14	3,15	1,92
Q3.6	3,27	2,07
Q3.16	3,30	2,04
Q3.11	3,37	2,03
Q3.23	3,40	1,43
Q3.2	3,45	2,18
Q3.7	3,47	2,06
Q3.5	4,02	1,96
Q3.21	4,15	1,61
Q3.4	4,15	2,05
Q3.1	4,22	2,11

Q3.22	4,31	1,41
Q3.15	4,32	1,81
Q3.13	4,44	1,94
Q3.24	4,87	1,81
Q3.20	5,32	,95
Moyenne	3,40	

*numéro de l'item en gras = différence significative entre les ensembles (Q3.12, Q3.6, Q3.2, Q3.5, Q3.4, Q3.1, Q3.13)

Tableau 5.5 Moyenne et écart type pour chaque ensemble (*Instruments de mesure*)

Ensemble 1			Ensemble 2			Ensemble 3		
Item	Moyenne	Écart type	Item	Moyenne	Écart type	Item	Moyenne	Écart type
Q3.12	1,29	0,75	Q3.12	1,40	1,05	Q3.18	1,46	1,04
Q3.18	1,72	1,47	Q3.18	1,51	1,16	Q3.19	2,04	1,55
Q3.10	1,76	1,32	Q3.10	1,60	1,29	Q3.12	2,14	1,65
Q3.2	2,34	1,87	Q3.8	1,98	1,55	Q3.17	2,18	1,61
Q3.9	2,41	1,74	Q3.19	2,53	1,77	Q3.10	2,29	1,78
Q3.3	2,71	2,03	Q3.1	2,76	2,12	Q3.8	2,46	1,92
Q3.6	2,78	2,00	Q3.6	2,98	1,99	Q3.9	2,68	1,98
Q3.8	2,78	1,85	Q3.9	2,98	1,97	Q3.14	2,82	1,77
Q3.19	2,83	1,92	Q3.3	3,00	1,97	Q3.16	2,86	1,94
Q3.17	2,97	1,90	Q3.14	3,13	1,87	Q3.2	2,93	2,02
Q3.11	2,97	2,05	Q3.17	3,20	1,79	Q3.3	3,07	2,09
Q3.16	3,28	2,13	Q3.7	3,22	2,03	Q3.11	3,11	2,06
Q3.14	3,31	2,05	Q3.23	3,47	1,41	Q3.13	3,39	2,01
Q3.23	3,31	1,45	Q3.16	3,60	1,96	Q3.23	3,46	1,45
Q3.7	3,48	2,05	Q3.4	4,02	2,07	Q3.15	3,79	1,95
Q3.5	3,55	2,10	Q3.21	4,02	1,62	Q3.7	3,86	2,12
Q3.4	3,71	2,17	Q3.11	4,07	1,84	Q3.5	4,29	1,80
Q3.22	4,16	1,45	Q3.22	4,31	1,41	Q3.21	4,29	1,44
Q3.21	4,17	1,70	Q3.24	4,44	2,11	Q3.22	4,61	1,32
Q3.15	4,33	1,89	Q3.5	4,47	1,78	Q3.6	4,79	1,64
Q3.13	4,81	1,81	Q3.13	4,62	1,87	Q3.1	4,93	1,65
Q3.1	5,02	1,68	Q3.15	4,64	1,57	Q3.24	5,07	1,51
Q3.24	5,10	1,65	Q3.2	5,20	1,46	Q3.20	5,25	1,08
Q3.20	5,28	1,02	Q3.20	5,42	0,78	Q3.4	5,25	1,27
Moyenne	3,34			3,44			3,46	

5.2.1 Instruments de mesure les moins utilisés

Item relatif aux entrevues

- Q3.12 Pour mon évaluation, je fais des entrevues (moyenne 1,51 et écart-type de 1,14).

Même s'il y a une différence statistiquement significative entre les moyennes concernant l'utilisation d'entrevues (Q3.12) comme instrument de mesure, elle n'est cependant pas populaire dans les trois ensembles, même si on l'utilise un peu plus dans l'ensemble 3.

Item relatif aux examens théoriques

- Q3.18 Dans mes examens théoriques, j'utilise des questions au choix (ex. : répondre à 2 questions sur 4 proposées) (moyenne 1,60 et écart-type de 1,28).

Dans les examens théoriques (Q3.18), on n'utilise pratiquement jamais de question de type « choix de répondre ». Par exemple : répondre à deux questions sur quatre.

Item relatif aux exposés oraux

- Q3.10 Pour mon évaluation, j'utilise des exposés oraux (moyenne 1,82 et écart-type de 1,43).

Il semble qu'on n'utilise presque jamais les exposés oraux (Q3.10) dans les ensembles 1 et 2. Ils semblent un peu plus utilisés dans l'ensemble 3 (moyenne 2,29, écart type 1,78).

5.2.1 Instruments de mesure les plus utilisés

Il y a plus de variabilité pour les items les plus élevés (Q3.20, Q3.24 et autres items selon l'ensemble) que pour les items faibles. On note aussi qu'il y a des items qui ont un écart type très élevé (en gras). Cela dénote une grande variabilité dans les réponses, ainsi la moyenne n'est pas très représentative. Dans un tel cas, nous avons plutôt à faire à deux groupes avec des modalités différentes pour une même pratique.

Items relatifs à l'utilisation du carnet de bord

- Q3.4 Pour mon évaluation, mes élèves produisent un carnet de bord pour décrire leur travail quotidien ou hebdomadaire (moyenne 4,15 et écart-type de 2,05).

La production d'un carnet de bord pour décrire leur travail quotidien ou hebdomadaire est populaire dans les cours de l'ensemble 2, mais avec un écart type important (2,07). Elle est encore plus populaire dans les cours de l'ensemble 3, mais pour cet ensemble, l'écart type est moins important ($s=1,27$). La figure 5.1 illustre la différence entre les trois ensembles par rapport à l'utilisation du carnet de bord. Cet instrument d'évaluation est nettement plus utilisé dans l'ensemble 3 comparativement aux deux autres ensembles.

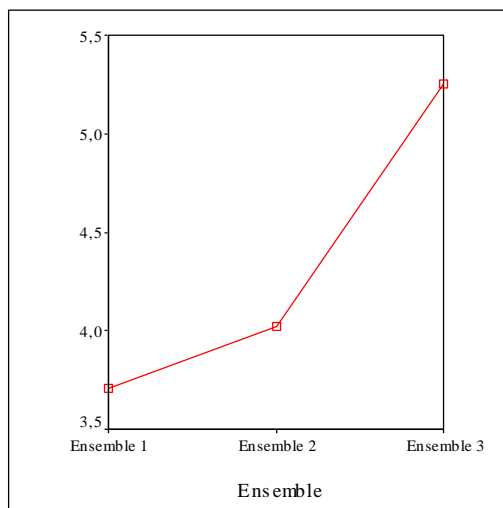


Figure 5.1 Illustration des moyennes de l'item Q3.4 pour chaque ensemble.

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q3.4 (1-3 et 2-3)	3,71	2,17	4,02	2,07	5,25	1,27

Items relatifs aux tests de condition physique

- Q3.1 J'utilise des tests de condition physique dans mes évaluations (moyenne 4,22 et écart-type de 2,11).

Les éducateurs physiques des cégeps utilisent beaucoup les tests de condition physique dans l'ensemble 1 (moyenne de 5,02) et l'ensemble 3 (moyenne de 4,93) pour évaluer la compétence des étudiants (c.-à-d. : critères de performance ou éléments de compétence). Si les tests de condition physique sont également utilisés dans l'ensemble 2, c'est dans une proportion beaucoup moins importante (moyenne de 2,76).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q3.1 (1-2 et 2-3)	5,02	1,68	2,76	2,12	4,93	1,65

Items relatifs aux tâches d'intégration

- Q3.22 J'évalue la compétence avec des tâches d'intégration (tâches sollicitant tous les éléments de la compétence) (moyenne 4,31 et écart-type de 1,41).
- Q6.21 Les tâches d'intégration (tâches sollicitant tous les éléments de la compétence) qui servent à l'évaluation sont inscrites dans mon plan de cours (moyenne 4,69 et écart-type de 1,57).

Pour évaluer la compétence, on semble utiliser des tâches d'intégration qui sollicitent plusieurs éléments de la compétence et ces tâches d'intégration sont inscrites dans le plan de cours (voir section Qualité de l'évaluation : question 6.21).

Items relatifs aux examens théoriques écrits.

- Q3.13 Pour mon évaluation, je fais des examens théoriques écrits (moyenne 4,44 et écart-type de 1,94).
- Q3.15 Dans mes examens théoriques, j'utilise des questions à réponse courte (moyenne 4,32 et écart-type de 1,81).

On utilise beaucoup les examens théoriques dans les ensembles 1 et 2 et moins souvent dans l'ensemble 3. Dans ces examens théoriques, on semble utiliser le plus souvent des questions à réponse courte.

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q3.13 (1-3 et 2-3)	4,81	1,81	4,62	1,87	3,39	2,01

Items relatifs aux l'épreuve synthèse de cours.

- Q3.24 Je fais une épreuve-synthèse de cours (épreuve finale) (moyenne 4,87 et écart-type de 1,81).

Selon nos statistiques descriptives, près de 75 % de nos répondants affirment faire des examens synthèses de cours (épreuve finale) souvent et toujours. Seulement 13 % n'en font jamais. C'est une pratique qui est courante dans le réseau et qu'on semble encourager dans les PIEA des collèges.

Items relatifs au le lien entre l'instrument et la compétence.

- Q3.20 Lorsque j'utilise un instrument de mesure théorique ou pratique, je m'assure qu'il est en lien direct avec la compétence (moyenne 5,32 et écart-type de 0,95).

Que ce soit pour évaluer la théorie ou la pratique, les enseignants en éducation physique s'assurent très majoritairement que leurs instruments de mesure sont en lien direct avec la compétence. Toutefois, on ne sait pas si l'instrument est en lien avec une partie de la compétence ou avec la compétence dans sa globalité. L'analyse d'autres items du questionnaire nous porte à croire que les instruments évaluent seulement une partie de la compétence.

5.2.3 Instruments de mesure : Analyse FPO versus FPC

Pour le concept *Instruments de mesure*, nous avons regroupé deux ensembles d'items qui constituent des indices FPC et FPO, et nous avons calculé la moyenne de chacun des deux ensembles d'items.

Examinons la différence entre les items FPC et les items FPO de la section *Instruments de mesure*. Comme il n'y a que les items Q3.4 et Q3.13 qui ont une différence significative entre les 3 ensembles, nous analyserons les 3 ensembles regroupés.

On se base sur la nature des items pour les regrouper, mais la composition de ces indices est relativement arbitraire, bien que la littérature dans le domaine reconnaît généralement ces caractéristiques (Voir le tableau 5.6). De plus, on doit être conscient que ces indices ne constituent pas des mesures unidimensionnelles. Ces indices sont :

- Indice FPC : Q3.3, Q3.4, Q3.7, Q3.21, Q3.22, Q3.24.
- Indice FPO : Q3.8, Q3.9, Q3.13, Q3.15, Q3.16, Q3.17.

Tableau 5.6 Moyenne et écart type des indices FPC et FPO pour la section *Instruments de mesure*.

Variable	<i>n</i>	<i>M</i>	<i>s</i>
C2MoyFPC	131	3,97	0,94
C2MoyFPO	131	3,33	1,14

n = nombre de répondants; *M* = moyenne; *s* = écart type

Dans le tableau 5.6, on note qu'il y a, cette fois, une différence significative entre les deux regroupements, et elle est à l'avantage de la FPC ($F_{(1, 130)} = 21,1$; $p =$

,001). Ce résultat est surprenant si on le compare aux résultats antérieurs. Toutefois, c'est une réalité que les enseignants en éducation physique ont des instruments qui sont assez compatibles avec la formation par compétences, mais pas nécessairement basés sur des tâches d'intégration (tâches qui activent tous les éléments de la compétence). Le problème se situe donc davantage sur les modalités de l'évaluation et sur les instruments de jugement, ainsi que la notation, comme nous le verrons dans les sections suivantes. Les analyses de certains items nous amènent à croire qu'il y aurait une mauvaise compréhension de la nature d'une *tâche d'intégration* en formation par compétences, il serait donc pertinent de se questionner sur le sens qu'on donne à une tâche d'intégration (Voir en annexe 2, un exemple de tâche d'intégration, basée sur un schéma intégrateur de l'ensemble 1 (Annexe 1)).

5.3 Instruments de jugement

Les instruments de jugement diffèrent des instruments de mesure, car ils ne servent pas à recueillir des informations, mais à porter un jugement à partir des instruments de mesure (*Exemple : je recueille des informations à partir d'une tâche ou d'un portfolio, mais je note ou je cote à partir d'une grille d'évaluation ou une liste de critères d'évaluation. Même pour un examen traditionnel, composé d'une série de questions déjà pondérées, je peux construire un instrument de jugement pour noter chaque réponse en fonction d'un résultat attendu.*) Dans ce questionnaire, les instruments de jugement concernent les critères de performance, les critères d'évaluation, les éléments de la compétence, les différents types de grilles, dont les grilles descriptives, avec des échelons correspondant à différents niveaux de développement de la compétence évaluée.

Le tableau 5.7 présente la moyenne et l'écart type pour les items de la section *Instruments de jugement*.

Tableau 5.7 Moyenne et écart type des items de la section *Instruments de jugement* (3 ensembles regroupées)

Item	Moyenne	Écart type
Q4.9	1,61	1,29
Q4.3	2,35	1,47
Q4.5	2,86	1,76
Q4.8	3,14	1,70
Q4.1	3,35	1,80
Q4.10	3,70	1,97
Q4.7	3,90	1,70
Q4.6	4,00	1,84
Q4.4	4,38	1,54
Q4.2	4,43	1,50

Note. Les numéros des items en gras indiquent une différence significative entre les ensembles (Q4.6-Q4.8). C'est essentiellement l'ensemble 2 qui est différent des 2 autres ensembles. Cela semble cohérent.

5.3.1 Instruments de mesure les moins utilisés

Item relatif à la notation à partir d'une cote

- Q4.9 Je note à partir d'une cote (ex. : A, B, C, D, E) (Moyenne 1,61, écart-type 1,29).

L'utilisation d'une cote (ex. : A, B, C, D, E) pour noter les étudiants est très rare. 76 % n'utilisent jamais cette forme de notation et seulement 2 % toujours.

Items relatifs à l'utilisation de critères de performance autres que ceux du devis.

- Q4.3 Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise des critères autres que les critères de performance de la compétence du devis (Moyenne 2,35, écart-type 1,48).

- Q4.5 Pour porter un jugement sur le degré de réussite de l'étudiant, je rédige moi-même les critères d'évaluation de la compétence (Moyenne 2,86, écart-type 1,76).

Les enseignants n'utilisent presque jamais ou très rarement des critères autres que les critères de performance de la compétence du devis ou des critères qu'ils ont rédigés eux-mêmes pour porter un jugement sur les apprentissages de l'étudiant.

5.3.2 Instruments de mesure les plus utilisés

Il n'y a pas vraiment d'items forts parmi les items relatifs aux *instruments de jugement*.

Items relatifs aux critères d'évaluation et de performances pour porter un jugement sur le degré de réussite de l'étudiant

- **Q4.6** Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise des critères d'évaluation sélectionnés par le département (Moyenne 4,00, écart-type 1,84).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q4.6 (1-2)	4,5	1,7	3,5	1,9	3,8	1,9

• Q4.4 Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise les éléments de la compétence du devis comme critères d'évaluation de la compétence (Moyenne 4,38, écart-type 1,54).

• Q4.2 Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise tous les critères de performance de la compétence du devis (Moyenne 4,43, écart-type 1,50).

Pour porter un jugement sur le degré de réussite de l'étudiant, les enseignants utilisent les critères d'évaluation sélectionnés par le département. C'est dans l'ensemble 1 qu'on note la moyenne d'utilisation la plus élevée et statistiquement différente avec les autres ensembles.

Même si les moyennes ne sont pas très fortes, ils utilisent également, soit les éléments de la compétence du devis, ou soit tous les critères de performance du devis pour porter un jugement sur le degré de réussite de l'étudiant. Nous devons toutefois relever une possible incohérence dans ces résultats. En effet, il est très curieux que les enseignants puissent, à la fois, utiliser les éléments de la compétence et les critères de performance pour leurs évaluations

5.3.3 Instrument de jugement : Analyse FPO versus FPC

Pour le concept *Instruments de jugement*, il est impossible de créer des indices FPC et FPO, car les items ne sont pas suffisamment spécifiques à l'une ou l'autre des approches. On devra donc analyser les items en fonction de leur moyenne et de leur écart type, ainsi que la nature du libellé (Voir tableau 4.1).

On note d'abord qu'il n'y a pas d'items qui ont une moyenne très élevée. Par contre, on observe des moyennes très basses (Q4.9- Q4.3- Q4.5). Les items ayant les moyennes les plus élevées portent à croire que la formation par compétences est très présente dans l'évaluation. Mais, des résultats aux items Q4.7 et Q4.8 (utilisation de grilles descriptives) montrent qu'il y a probablement une mauvaise compréhension de la FPC. Par exemple, l'item Q4.7 (utilisation d'une grille descriptive) obtient une moyenne de 3,9 sur 6, alors que les présentations des modes d'évaluation, lors des colloques en éducation physique montrent clairement qu'un tel type de grille n'est pratiquement jamais utilisé (Voir un exemple de grille descriptive à échelons, en Annexe 3). Il est donc très difficile d'interpréter les résultats de cette section (Instruments de jugement). C'est à notre avis, la faiblesse majeure de l'évaluation en éducation physique (avec celle de la rédaction des compétences 1 et 3 (devis ministériels)). Nous commençons à penser sérieusement qu'il faudrait réaliser une recherche qualitative, avec des entrevues, afin de mieux saisir les problèmes liés à l'évaluation par compétences. De plus, il devient évident qu'une révision en profondeur des devis ministériels serait nécessaire pour rendre l'évaluation cohérente avec le sens qu'on veut donner aux différentes compétences en éducation physique. Évidemment, si on veut arriver à bien opérationnaliser l'évaluation, il est impérieux d'arriver à rédiger un guide d'évaluation qui oriente mieux les éducateurs physiques, particulièrement quant à l'utilisation des critères de performance, des critères d'évaluation de la compétence et à l'intégration des éléments de la compétence dans des tâches d'intégration.

5.4 Composition de la note

Le concept *Composition de la note* réfère aux différentes façons d'accorder des points ou des cotes en vue d'une évaluation sommative ou certificative. Avec ce concept, on tente d'identifier avec précision les objets sur lesquels on donne des notes et dans quelle proportion par rapport à la note finale. Le concept couvre aussi le ou les évaluateurs potentiels (l'enseignant, l'étudiant, les pairs, l'équipe de travail). La composition de la note réfère aussi certains aspects particuliers, tels l'évaluation du progrès, des performances, les absences, les attitudes, les comportements et la qualité de la langue.

Le tableau 5.8 présente la moyenne et l'écart type de chacun des items associés au concept *Composition de la note*. Dans cette section, il y a suffisamment d'items avec une différence significative entre les 3 ensembles pour justifier une présentation par ensemble (items avec une différence significative : Q5.1- Q5.19- Q5.13- Q5.15- Q5.16- Q5.17. Le tableau 5.9 présente la moyenne et l'écart type des items de cette section.

Tableau 5.8 Moyenne et écart type des items de la section *Composition de la note* (3 ensembles regroupés)

Item	Moyenne	Écart-type
Q5.14	1,46	1,05
Q5.2	1,47	1,21
Q5.1	1,58	1,38
Q5.15	1,63	1,31
Q5.20	1,65	1,32
Q5.10	2,17	1,91
Q5.16	2,37	1,74
Q5.13	2,51	1,93
Q5.3	2,57	1,96
Q5.17	2,64	1,90
Q5.25	2,71	2,14
Q5.11	2,81	2,19
Q5.8	3,11	1,86
Q5.23	3,11	2,07
Q5.18	3,21	1,95
Q5.24	3,33	2,00

Q5.22	3,72	2,03
Q5.27	3,76	1,86
Q5.9	4,02	1,75
Q5.21	4,09	2,13
Q5.7	4,17	1,79
Q5.4	4,31	1,63
Q5.26	4,42	1,84
Q5.6	4,44	1,58
Q5.19	4,47	1,95
Q5.5	4,56	1,56
Q512	4,70	1,91

Tableau 5.9 Moyenne et écart type des items pour chaque ensemble (Composition de la note)

Ensemble 1			Ensemble 2			Ensemble 3		
Item	Moyenne	Écart type	Item	Moyenne	Écart type	Item	Moyenne	Écart type
Q5.1	1,17	,68	Q5.2	1,56	1,32	Q5.2	1,50	1,26
Q5.14	1,24	,71	Q5.20	1,69	1,41	Q5.14	1,50	1,23
Q5.15	1,34	,95	Q5.14	1,71	1,24	Q5.15	1,54	1,37
Q5.2	1,38	1,11	Q5.10	1,89	1,70	Q5.20	1,68	1,33
Q5.20	1,60	1,27	Q5.1	2,00	1,68	Q5.1	1,75	1,71
Q5.16	1,98	1,50	Q5.15	2,04	1,58	Q5.10	1,75	1,48
Q5.13	2,07	1,74	Q5.13	2,42	1,82	Q5.11	2,18	1,74
Q5.17	2,21	1,68	Q5.25	2,62	2,05	Q5.25	2,21	2,01
Q5.3	2,26	1,78	Q5.3	2,67	2,00	Q5.16	2,36	1,79
Q5.10	2,59	2,18	Q5.11	2,71	2,18	Q5.23	2,79	1,97
Q5.8	2,83	1,83	Q5.16	2,89	1,90	Q5.17	2,86	2,09
Q5.25	3,02	2,25	Q5.18	2,93	1,78	Q5.3	3,07	2,19
Q5.11	3,19	2,34	Q5.23	2,96	2,07	Q5.8	3,18	1,81
Q5.24	3,21	2,04	Q5.17	3,07	1,96	Q5.18	3,25	2,08
Q5.23	3,40	2,12	Q5.27	3,40	1,78	Q5.24	3,25	1,96
Q5.18	3,40	2,02	Q5.8	3,44	1,91	Q5.22	3,50	1,99
Q5.9	3,60	1,86	Q5.24	3,53	2,01	Q5.13	3,57	2,15
Q5.22	3,66	2,08	Q5.22	3,93	2,02	Q5.27	3,93	1,78
Q5.21	3,76	2,21	Q5.7	4,00	1,76	Q5.4	4,14	1,63
Q5.27	3,97	1,95	Q5.26	4,27	1,88	Q5.21	4,14	2,14
Q5.7	4,12	1,90	Q5.9	4,27	1,75	Q5.19	4,14	2,14
Q5.4	4,24	1,77	Q5.6	4,33	1,57	Q5.5	4,36	1,59
Q5.6	4,40	1,71	Q5.19	4,40	1,95	Q5.9	4,50	1,32
Q5.26	4,45	1,85	Q5.4	4,49	1,44	Q5.7	4,54	1,57
Q512	4,52	2,00	Q5.21	4,49	2,00	Q5.26	4,61	1,79
Q5.5	4,67	1,63	Q5.5	4,53	1,49	Q5.6	4,68	1,33
Q5.19	4,67	1,88	Q5.12	4,73	1,57	Q5.12	5,04	1,82

5.4.1 Composantes de la note les moins utilisées

Items relatifs à l'évaluation par les pairs

- Q5.14 Une partie des points de l'évaluation provient de l'évaluation par les pairs (moyenne 1,46 et écart type 1.05)

- Q5.15 Une partie des points de l'évaluation provient de l'évaluation par les pairs, combinée à la mienne (moyenne 1,63 et écart type 1,31).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q5.15 (1-2)	1,3	0,9	2,0	1,6	1,5	1,4

On n'utilise pas beaucoup l'évaluation par les pairs, même si l'évaluation des pairs peut être combinée à celle de l'enseignant. C'est dans l'ensemble 2 que cette pratique est la plus utilisée, mais rarement (18 %). On peut considérer la faible utilisation de l'évaluation par les pairs comme une bonne pratique en fonction de la FPC.

Items relatifs aux points pour le progrès

- Q5.20 Je donne des points pour le progrès, sans utiliser un prétest et un post-test (moyenne 1,65 et écart type 1.32).

Très peu d'enseignants attribuent des points pour l'amélioration sans utiliser un prétest en début de session et un post-test vers la fin.

Items relatifs à l'attribution de points lors de tournois compétitifs.

- Q5.2 J'accorde des points pour les résultats d'équipe dans des tournois compétitifs (ex. : 2 points pour une victoire (moyenne 1,47 et écart type 1,21).

- Q5.1 J'accorde des points pour les résultats individuels dans des tournois compétitifs (moyenne 1,58 et écart type 1,38).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q5.1 (1-2)	1,2	0,7	2,0	1,7	1,8	1,7

Les enseignants n'accordent à peu près pas de points pour les résultats d'équipe ou individuels dans des tournois compétitifs.

5.4.2 Composantes de la note les plus utilisées

Item relatif aux points pour la langue écrite

- Q5.26 J'enlève des points pour un usage inapproprié de la langue écrite (moyenne 4,42 écart type 1,84).

La majorité des enseignants enlèvent des points pour un usage inapproprié de la langue écrite. Cependant 26 % d'entre eux en enlèvent rarement, très rarement ou jamais de points pour la langue écrite.

Item relatif aux absences aux cours

- Q5.12 Trop d'absences entraîne automatiquement un échec (moyenne 4,70 écart type 1,91).

Une grande majorité d'enseignants attribuent un échec à l'étudiant lorsqu'il manque trop de cours. Par ailleurs, on sait que dans plusieurs cégeps, il y a possibilité de reprendre une ou des absences selon diverses considérations.

Item relatif à la notation de l'épreuve finale

- Q5.6 J'utilise les critères de performance de la compétence pour noter l'épreuve finale (moyenne 4,44, écart type 1,58).

On semble surtout utiliser les critères de performance, et non les éléments de compétence pour noter l'épreuve finale. Cependant un total de 21 % des répondants ne les utilisent jamais (11 %), très rarement (3 %), et rarement (7 %).

Item relatif à la responsabilité de noter

- Q5.19 Je suis le seul à donner des notes à mes étudiants (moyenne 4,47, écart type 1,95).

Même si la grande majorité des enseignants affirment être les seuls à donner des notes à leurs étudiants, on peut s'étonner qu'un total de 30 % semblent partager cette responsabilité, et ce, particulièrement dans l'ensemble 3 (45 %).

Item relatif aux points attribués aux éléments de la compétence

- Q5.5 J'attribue des points à chaque élément de la compétence (moyenne 4,56 écart type 1,56).

Même si la majorité des enseignants semblent attribuer des points à tous les éléments de la compétence, il n'en demeure pas moins que près de 10 % n'attribuent pas de points à tous les éléments.

Item relatif aux points pour le progrès

- Q5.21 Je donne des points pour le progrès, en utilisant un prétest et un post-test (moyenne 4,09 et écart type 2,13).

Item	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q5.21	3,76	2,21	4,49	2,00	4,14	2,14

Beaucoup d'enseignants accordent des points pour le progrès dans les trois ensembles. Cependant, des écarts types très importants ($s > 2$) pour tous les ensembles nous indiquent des variations importantes qu'en a l'attribution de points pour le progrès en utilisant un pré test et un post-test. Ces résultats semblent cohérents avec ceux observés à la question Q5.20 et nous indique que lorsqu'on attribut des points pour le progrès, c'est en vérifiant l'amélioration entre le début et la fin de la session.

- Q5.20 Je donne des points pour le progrès, sans utiliser un prétest et un post-test (moyenne 1,65 et écart type 1.32).

5.4.3 Composition de la note : Analyse FPO versus FPC

Bonnes pratiques versus mauvaises pratiques

Nous avons tenté de composer des indices en fonction de la FPC et de la FPO (Composition de la note), mais même si on obtient des coefficients alpha raisonnables, il est impossible de trouver un lien logique dans ce regroupement d'items. Nous avons donc composé des indices de bonnes pratiques et de mauvaises pratiques.

- indice bonnes pratiques : Q5.5- Q5.7- Q5.9- Q5.19- Q5.21-

- indice mauvaises pratiques : Q5.1- Q5.2- Q5.10- Q5.14- Q5.15- Q5.16- Q5.17- Q5.18- Q5.20

Tableau 5.10 Moyenne et écart type des indices de *bonnes pratiques* et de *mauvaises pratiques* pour la section *Composition de la note*.

Variable	<i>n</i>	<i>M</i>	<i>s</i>
C4 bonnes	131	4,26	,93
C4 mauvaises	131	2,02	,76

n = nombre de répondants; *M* = moyenne; *s* = écart type

Il y a une différence significative en faveur des bonnes pratiques ($F_{(1, 130)} = 401,2$; $p = ,001$). Toutefois, ces bonnes pratiques ne sont pas toutes reliées à la FPC. De plus, plusieurs mauvaises pratiques obtiennent des moyennes relativement élevées, par exemple : Q5.10 (présences aux cours), Q5.16 (auto-évaluation), Q5.17 (coévaluation), Q5.18 (effort).

Et que dire de Q5.27 (Critère de performance : communication claire et argumentée), avec une moyenne de 3,76; comment a-t-on évalué ce critère? Oui, il y aurait lieu de conduire une recherche qualitative pour comprendre ces résultats bizarres.

5.5 Qualités des évaluations

Ce concept concerne la concordance entre les objets d'évaluation, le plan de cours et la nature des évaluations. Le questionnaire considère autant les instruments de mesure que les instruments de jugement. Le concept se rapporte aux qualités généralement exigées en évaluation, telles la rigueur, la validité, la clarté, la discrimination, la standardisation. Le questionnaire vise à identifier les biais possibles en évaluation, tels l'influence de la force de la classe, le jugement des collègues, la propension ou l'incapacité à faire échouer un étudiant, le comportement d'un étudiant, la distribution des notes en relation avec la courbe normale.

Il y a une différence significative entre les 3 ensembles sur seulement 2 items (Q6.1 et Q6.17). Nous présentons donc les items pour les 3 ensembles regroupés dans le tableau 5.11.

Tableau 5.11 Moyenne et écart type des items pour la section *Qualités des évaluations*.

Item	Moyenne	Écart type
Q6.13	1,80	1,27
Q6.12	1,84	1,26
Q6.9	2,06	1,49
Q6.6	2,24	1,52
Q6.16	2,64	1,50
Q6.20	2,84	1,50
Q6.15	2,92	1,39
Q6.14	3,27	1,36
Q6.17	3,90	1,28
Q6.19	3,90	1,54
Q6.18	4,03	1,54
Q6.10	4,07	1,87
Q6.8	4,11	1,69
Q6.23	4,69	1,22
Q6.21	4,69	1,57
Q6.22	4,95	1,21
Q6.3	4,95	1,54
Q6.5	5,13	1,08
Q6.24	5,31	,90
Q6.4	5,38	,71
Q6.1	5,56	,79

Q6.11	5,64	,90
Q6.7	5,66	,75
Q6.2	5,79	,55

Numéro en gras = différence significative entre les ensembles.

5.5.1 Qualités des évaluations les moins fréquentes

Item relatif aux résultats passés de l'étudiant.

- Q6.13 Les résultats passés d'un étudiant influencent ses futures évaluations (moyenne 1,80 écart type 1,27).

Le résultat est très clair, les évaluations des enseignants en éducation physique ne sont pas influencées par les résultats passés d'un étudiant. Le fait que l'enseignant n'a pas directement accès à ces résultats n'est peut être pas étranger à cette situation.

Item relatif à la force de la classe.

- Q6.12 La force de la classe influence mes évaluations (moyenne 1,84 écart type 1,26).

La force de la classe, ne semble pas influencer les évaluations des enseignants.

Item relatif à la comparaison des étudiants entre eux

- Q6.9 Mes évaluations visent à comparer les étudiants entre eux (évaluations normatives) (moyenne 2,06 écart type 1,49).

5.5.2 Qualités des évaluations les plus fréquentes

Item relatif aux instruments évaluant les objets d'évaluation

- **Q6.1** Mes instruments évaluent précisément les objets d'évaluation prévus dans le plan de cours (moyenne 5,57 écart type 0,79).

	Ensemble 1		Ensemble 2		Ensemble 3	
	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>	<i>M</i>	<i>s</i>
Q6.1 (2-3)	5,6	0,7	5,3	1,0	5,8	0,5

La moyenne très élevée et l'écart type très faible à cette question nous indiquent que les enseignants ont vraiment l'impression d'évaluer précisément les objets d'évaluation prévus dans le plan de cours.

Item relatif à l'équité

Q6.11 J'ai un souci d'équité dans mes évaluations (moyenne 5,64 écart type ,90).

La moyenne très élevée et l'écart type très faible à cette question nous indiquent que l'équité est une qualité de l'évaluation très importante pour les enseignants en éducation physique.

Item relatif aux consignes d'évaluation.

- **Q6.7** Mes consignes d'évaluation sont connues de l'étudiant avant l'évaluation (moyenne 5,66 écart type ,75).

Les enseignants en éducation physique ont la presque certitude (94 %) que les étudiants ont été mis au courant des consignes d'évaluation avant l'évaluation.

Item relatif à la nature de mes évaluations.

- | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Q6.2 La nature de mes évaluations est décrite dans le plan de cours (ex. : travaux, examens, évaluations pratiques) (moyenne 5,79 écart type ,55). |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Selon les enseignants en éducation physique, la nature des évaluations du cours est décrite dans le plan de cours distribué à l'étudiant en début de session.

5.5.3 Qualités des évaluations : Analyse FPO versus FPC

Nous avons tenté de composer des indices en fonction de la FPC et de la FPO (Qualités des évaluations), mais il a été impossible de trouver un lien logique dans ces regroupements d'items. Nous avons donc composé des indices de bonnes pratiques et de mauvaises pratiques.

Pour cette section, on remarque que ce sont surtout les bonnes pratiques qui présentent des moyennes élevées et les mauvaises pratiques qui ont des moyennes basses. C'est très encourageant de constater que les problèmes en évaluation ne viennent pas des qualités généralement reconnues en évaluation, mais plutôt de la non-application des principes d'évaluation propres à la formation par compétences (référons-nous aux modalités d'évaluation, aux instruments de jugement et à la composition de la note : plusieurs pratiques reliées à la FPO obtiennent des moyennes élevées et plusieurs pratiques associées à la FPC présentent des moyennes basses). De plus, dans ces sections, plusieurs items présentent des écarts types très élevés, ce qui dénote que les enseignants sont très partagés par rapport à certaines pratiques.

Étant donné que nous n'avons pas réussi à créer des indices propres à la FPC et à la FPO, nous avons composé un indice bonnes pratiques versus un indice

mauvaises pratiques afin de vérifier plus précisément ce qui s'observe en examinant un à un les items.

- indice bonnes pratiques : Q6.1-Q6.2-Q6.3-Q6.5-Q6.8-Q6.10-Q6.11-Q6.18-Q6.19-Q6.21-Q6.23
- indice mauvaises pratiques : Q6.6- Q6.9- Q6.12- Q6.13- Q6.14- Q6.16- Q6.17

Le tableau suivant 5.12 présente les résultats par rapport aux bonnes pratiques et aux mauvaises pratiques en évaluation. Les bonnes pratiques l'emportent grandement sur les mauvaises pratiques. Toutefois, certaines mauvaises pratiques présentent des moyennes relativement élevées : Q6.14 (accorder une note de passage si près de 60 %), Q6.16 (influence du jugement de mes collègues), Q6.17 (distribution des notes suit la courbe normale,)

Tableau 5.12 Moyenne et écart type des indices *bonnes pratiques* et *mauvaises pratiques* pour la section *Qualités des évaluations*.

Variable	<i>n</i>	<i>M</i>	<i>s</i>
C5bonnes	131	4,78	,63
C5mauvaises	131	2,54	,68

n = nombre de répondants; *M* = moyenne; *s* = écart type

6. CONCLUSION

Le tableau 6.1 présente, pour chaque dimension des pratiques évaluatives une synthèse montrant quelle approche est dominante en fonction de la FPC ou de la FPO. Le tableau indique par le signe « + » ou « - » l'approche dominante. La catégorisation selon la FPC ou la FPO est basée sur l'analyse des moyennes des indices créés pour chacune des deux approches et suivie d'une analyse de la variance (ANOVA, mesures répétées) pour établir s'il y avait une différence significative entre les deux approches (FPC ou FPO). Toutefois, pour deux dimensions, nous avons dû analyser les items en fonction d'indices relatifs aux bonnes pratiques ou aux mauvaises pratiques en évaluation. Il faut noter que les bonnes pratiques ne sont pas exclusives à la FPC. Le jugement pour ces dimensions est donc moins précis, mais il a été tout de même possible de catégoriser les dimensions selon qu'elles étaient davantage axées sur la FPC ou la FPO.

Les résultats présentés dans le tableau 6.1 montrent une tendance assez forte vers des pratiques évaluatives orientées vers la FPO. En effet, sur cinq dimensions, la FPO obtient 3 « + » (plus orientée) et 1 « - » (moins orientée), tandis que la FPC obtient 1 « + » (plus orientée) et 3 « - » (moins orientée). La dimension *Qualités des évaluations* n'a pu être catégorisée.

Tableau 6.1 Tableau synthèse du degré de la proximité des dimensions de l'évaluation en fonction de la FPC ou de la FPO.

DIMENSIONS	FPC	FPO
Modalités des évaluations	(-)	(+)
Instruments de mesure	(+)	(-)
Instruments de jugement	(-)	(+)
Composition de la note	(-)	(+)
Qualités des évaluations	(?)	(?)
TOTAL	3 – et 1 +	1 – et 3 +

Modalités des évaluations

Pour la dimension *Modalités des évaluations*, les évaluations sont plus orientées vers la FPO. Les principales pratiques évaluatives qui jouent en faveur de la FPC ou de la FPO sont les suivantes :

- plusieurs petites évaluations
- plusieurs évaluations comptant pour moins de 12 % de la note
- environ 50 % de professeurs n'ont pas d'épreuve synthèse comptant pour au moins 40 % de la note
- une majorité de professeurs donnent droit aux notes de cours pour les évaluations plutôt qu'exiger du par coeur

Instruments de mesure

Pour la dimension *Instruments de mesure*, les évaluations sont plus orientées vers la FPC. Les principales pratiques évaluatives qui jouent en faveur de la FPC ou de la FPO sont les suivantes :

- les enseignants en éducation physique ont des instruments qui sont assez compatibles avec la formation par compétences, mais pas nécessairement basés sur des tâches d'intégration
- les analyses de certains items nous amènent à croire qu'il y aurait une mauvaise compréhension de la nature d'une *tâche d'intégration* en formation par compétences (tâche qui active tous les éléments de la compétence)

Instruments de jugement

Pour la dimension *Instruments de jugement*, les évaluations sont plus orientées vers la FPO. Les principales pratiques évaluatives qui jouent en faveur de la FPC ou de la FPO sont les suivantes (notez que pour la dimension *Intruments de jugement*, il a été impossible de créer des indices FPC et FPO, nous avons donc créé des indices de bonnes pratiques et de mauvaises pratiques, toutefois les bonnes pratiques ne sont pas exclusives à la FPC) :

- Les items ayant les moyennes les plus élevées portent à croire que la formation par compétences est très présente dans l'évaluation. Mais, des résultats aux items relatifs aux grilles descriptives montrent qu'il y a probablement une mauvaise compréhension de la nature et de l'utilisation d'une grille d'évaluation descriptive qui sont propres à la FPC.

Composition de la note

Pour la dimension *Composition de la note*, les évaluations sont plus orientées vers la FPO. (Notez que pour la dimension *Composition de la note*, il a été impossible de créer des indices FPC et FPO, nous avons donc créé des indices de bonnes pratiques et de mauvaises pratiques.

- Il y a une différence significative en faveur des bonnes pratiques, toutefois, ces bonnes pratiques ne sont pas toutes reliées à la FPC. De plus, plusieurs mauvaises pratiques obtiennent des moyennes relativement élevées. À partir d'une analyse des items un à un et la mise en relation avec des items des autres dimensions, nous classons la dimension *Composition de la note* davantage axée sur la FPO.

Qualités des évaluations

Pour la dimension *Qualités des évaluations*, il a été impossible de créer des indices FPC et FPO, nous avons donc créé des indices de bonnes pratiques et de mauvaises pratiques. Même avec des indices bonnes pratiques et mauvaises pratiques, il a été impossible de catégoriser cette dimension dans la FPC ou la FPO.

- Les bonnes pratiques l'emportent grandement sur les mauvaises pratiques. Toutefois, certaines mauvaises pratiques présentent des moyennes relativement élevées.

Les problèmes en évaluation ne viennent pas des qualités généralement reconnues en évaluation, mais plutôt de la non-application des principes d'évaluation propres à la formation par compétences. En nous référant aux dimensions *Modalités des évaluations*, *Instruments de jugement* et à la *Composition de la note*, on note que des pratiques reliées à la FPO obtiennent des moyennes élevées et plusieurs pratiques associées à la FPC présentent des moyennes basses. De plus, dans ces sections,

plusieurs items présentent des écarts types très élevés, ce qui dénote que les enseignants sont très partagés par rapport à certaines pratiques.

L'analyse du tableau 6.1 et les remarques sur chacune des dimensions des pratiques évaluatives des professeurs d'éducation physique nous amènent à conclure que les éducateurs et les éducatrices en éducation physique ont plutôt de bonnes pratiques en évaluation, mais ils sont encore loin d'appliquer les principes d'évaluation propres à la formation par compétences. Il faut toutefois ajouter à cette conclusion, que les professeurs d'éducation physique ne sont ni plus, ni moins orientés sur la FPC que les professeurs des autres disciplines. Nous appuyons cette observation sur les résultats de l'étude réalisée par Bélanger et Tremblay (2012)¹, dans le cadre d'une collaboration du Collège Maisonneuve et du Regroupement des collèges PERFORMA.

La question de recherche de cette étude est la suivante : « Les croyances et les pratiques des enseignants de l'ordre collégial s'inscrivent-elles dans le prolongement de l'approche par compétences introduite par la réforme de l'enseignement collégial ? » L'échantillonnage de cette recherche par sondage fut constitué d'un échantillonnage aléatoire en grappes, sélectionné dans un premier niveau parmi les cégeps de la province, et dans un deuxième niveau, parmi les départements des cégeps sélectionnés. Cette recherche visait à analyser les croyances et les pratiques évaluatives sous cinq dimensions inspirées du modèle de Stiggins (1987, 2006)^{2 3} et adapté par Howe et Ménard (1993)⁴: but de l'évaluation, qualité de l'évaluation, intégrité du processus d'évaluation, rétroaction, note. La recherche de Bélanger et Tremblay (2012), en plus de dresser l'inventaire des pratiques évaluatives les plus

¹ Bélanger, D.-C., et Tremblay, K. (2012). *Portrait actualisé des croyances et des pratiques en évaluation des apprentissages au collégial*. Montréal : Collège Maisonneuve.

² Stiggins, R. (1987). NCME instructional Module on Design and Development of Performance Assessments. *Instructional Topics in Educational Measurement*, 33-42.

³ Stiggins, R. (2006). *Balanced Assessment Systems : Redefining Excellence in Assessment*. Educational Testing Service, Portland, 1-10

⁴ Howe, R. et Ménard, L. (1993). *Croyances et pratiques en évaluation des apprentissages : étude des croyances et des pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages*. Rapport de recherche PAREA, Collège Montmorency.

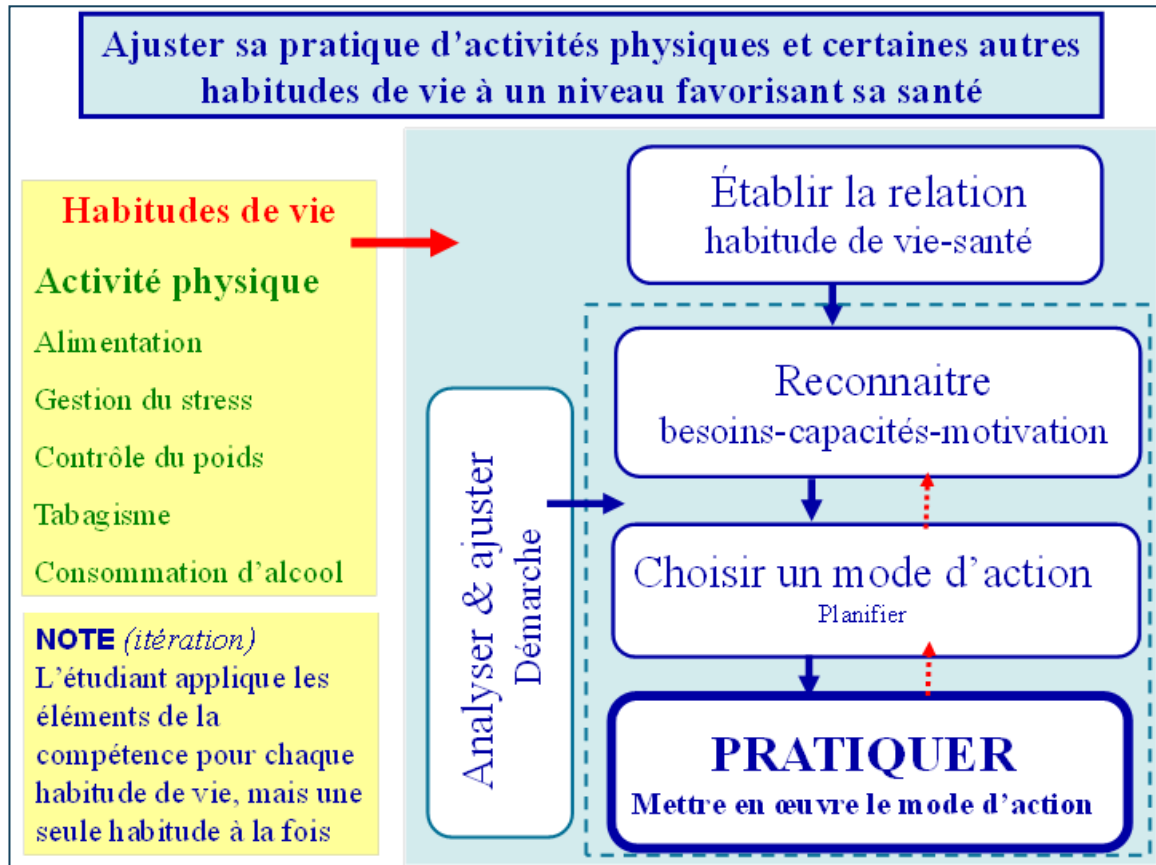
fréquentes et les moins fréquentes, analyse le degré de proximités des pratiques évaluatives des enseignants des cégeps avec l'approche par compétences. Selon cette étude, les pratiques en évaluation des apprentissages au collégial présentent une proximité assez élevée avec les principes généralement reconnus (PGR) en mesure et évaluation et cohérents avec l'approche par compétences (APC). Ce résultat est plutôt encourageant, toutefois il faut noter que, dans cette recherche, on a joint dans un même ensemble les pratiques généralement reconnues en mesure et évaluation et les pratiques propres à l'approche par compétences. Cela a contribué significativement à l'augmentation du degré de proximité comparativement à une mesure qui se serait limitée aux pratiques clairement associées à la formation par compétences. Les auteurs de cette recherche soulignent aussi que les pratiques en évaluation varient selon certaines caractéristiques, dont le secteur d'enseignement et la taille des cégeps. Globalement, la recherche révèle des changements marqués pour certaines pratiques (présence de l'évaluation formative, tâche complexe, épreuve finale, assez forte proportion des notes attribuées à l'épreuve finale, grille d'évaluation) et de grands pas qu'il reste à faire sur d'autres pratiques (répartition de notes, encore plusieurs petites évaluations, approche davantage centrée sur une évaluation sommative que sur une évaluation certificative).

ANNEXES

- **Annexe 1** Exemple d'un schéma intégrateur pour l'ensemble 1
- **Annexe 2** Exemple de tâche d'intégration pour l'ensemble 1
- **Annexe 3** Exemple d'une grille d'évaluation descriptive, avec échelons, pour l'ensemble 1
- **Annexe 4** Les questionnaires français et anglais

ANNEXE 1

Exemple d'un schéma intégrateur pour l'ensemble 1 (*F. Lasnier*)



Annexe 2

Exemple de tâche d'intégration pour l'ensemble 1 (F. Lasnier)

TÂCHE D'INTÉGRATION N° 1	
COMPÉTENCE : <i>Ajuster sa pratique d'activités physiques et certaines autres habitudes de vie à un niveau favorisant sa santé</i>	
SITUATION :	<i>(Contextualisation)</i>
<p>Au secondaire, vous avez développé la compétence « Adopter un mode de vie sain et actif » (<i>Vous vous en rappelez?</i>). En ce début de session au collégial, nous voulons vous sensibiliser au processus de la compétence « <i>Ajuster sa pratique d'activités physiques et certaines autres habitudes de vie à un niveau favorisant sa santé</i> ». Nous voulons donc approfondir votre capacité d'être autonome dans l'amélioration de vos habitudes de vie, en particulier, par rapport à la pratique de l'activité physique. Au cours de la session, nous aborderons différentes habitudes de vie (alimentation, gestion du stress, autres), mais notre première tâche d'intégration portera sur la pratique de l'activité physique. Cette tâche s'échelonnera sur toute la session, elle sera complétée par plusieurs activités d'apprentissage qui devront être conservées, car elles seront nécessaires pour réaliser votre rapport final ou votre épreuve synthèse de cours, afin de démontrer le degré de maîtrise de votre compétence développée dans ce cours. À la suite de ces activités d'apprentissage, nous procéderons à plusieurs évaluations formatives afin de vous familiariser avec le processus de la compétence (voir le schéma intégrateur) et s'assurer que vous avez bien compris l'application de notions théoriques nécessaires à la mise en oeuvre de votre compétence « <i>Ajuster sa pratique d'activités physiques et certaines autres habitudes de vie à un niveau favorisant sa santé</i> ».</p> <p>Cette tâche relative à la pratique de l'activité physique revêt une importance capitale, car de nombreuses recherches scientifiques montrent clairement que la pratique régulière et adéquate de l'activité physique constitue la meilleure protection contre les maladies cardiovasculaires, et même contre plusieurs types de cancer.</p> <p>Cette tâche d'intégration est très importante, car le processus appliqué servira de modèle pour les autres habitudes de vie qui seront abordées au cours de la session. (<i>Voir les consignes</i>). De plus, et cela n'est pas négligeable, l'acquisition du processus de cette compétence pourra vous servir tout au long de votre vie. La maîtrise de cette compétence vous assurera un fort degré d'autonomie pour réussir à pratiquer sainement diverses activités physiques.</p>	
TÂCHE :	
<p>➤ Produisez un rapport relatant l'ensemble de la démarche que vous avez mis en oeuvre pour ajuster votre pratique de l'activité physique à un niveau favorisant la santé.</p> <p style="text-align: center;"><i>----> (Résultat attendu, le produit à évaluer) Quoi</i></p>	
CONSIGNES :	
<ul style="list-style-type: none"> • Vous devez remettre votre rapport le 15 décembre. La présentation du rapport devra suivre les normes du département d'éducation physique. 	

- Vous devez : → (*E = élément de C.; A.A. = activité d'apprentissage*)
- (E1) Établir la relation entre votre pratique de l'activité physique et votre santé (voir A.A 1 : Recherches et notions théoriques; A.A. 2 : Mes comportements positifs et négatifs; A.A. 3 : Effets positifs et négatifs ressentis).
- (E2) Identifier vos besoins, vos capacités et vos motivations liés à la pratique de l'activité physique (voir A.A 4 : Résultats de mes évaluations de ma condition physique; A.A 5 : Relevé de mes besoins; A.A 6 : Identification de mes motivations personnelles).
- (E3) Présenter, de façon détaillée, en référence aux notions théoriques et en référence aux résultats de vos besoins identifiés, votre plan d'action (mode d'amélioration ou de maintien) pour chaque déterminant de la condition physique travaillé dans le cours (voir les A.A 6, 7, 8 et 9).
- (E4) Présenter la réalisation de votre plan d'action (A.A. 10 : Application des règles de sécurité; A.A. 11 : Fiches-synthèses de la pratique hebdomadaire de mes activités physiques; A.A. 12 : Réévaluation du mon plan d'action; A.A. 13 : Analyse critique de mon plan d'action et de sa réalisation)
- Utilisez la terminologie juste et usuelle en éducation physique et en conformité avec les notions scientifiques.

----> (Varient selon le niveau de développement, de spécifiques à globales) Comment

INTENTIONS :

- Se familiariser avec les éléments de la compétence 1.
- Acquérir des connaissances sur la santé et la pratique de l'activité physique.
- Faire les liens entre la tâche réalisée et les éléments de la compétence 1.
- Appliquer l'évaluation formative pour chaque activité d'apprentissage (A.A.)

----> Cibles d'apprentissage : compétence, contenu disciplinaire, stratégies

Annexe 3

Exemple d'une grille d'évaluation descriptive, avec échelons, pour l'ensemble 1 (F. Lasnier).

→ Veuillez lire les consignes pour l'utilisation de cette grille qui sont présentées à la fin de la grille.

ÉCHELONS ET INTERVALLES DE NOTES

Pour chaque critère

1	2	3	4
<i>Acquisition insuffisante</i>	<i>Acquisition minimale (seuil)</i>	<i>Acquisition attendue</i>	<i>Acquisition au-delà de l'attente</i>
0 À 59	60 À 69	70 À 89	90 À 100
- + _____	- + _____	- + _____	- + _____

GRILLE D'ÉVALUATION DESCRIPTIVE À ÉCHELONS

Compétence : Ajuster sa pratique d'activités physiques à un niveau favorisant sa santé			
Critère 1 - Relation entre son niveau de pratique d'activités physiques et sa santé			
1	2 (seuil de réussite)	3	4
Énonce seulement 1 lien scientifiquement démontré entre l'activité physique et la santé, ou présente des liens de façon floue. Énonce 1 comportement (positifs ou négatifs) affectant sa propre santé. Identifie nébuleusement l'effet ressenti physiologiquement pour chacun des comportements identifiés.	Énonce clairement 2 liens scientifiquement démontrés entre l'activité physique et la santé. Énonce clairement 2 comportements (positifs ou négatifs) affectant sa propre santé. Identifie clairement l'effet ressenti physiologiquement pour chacun des 2 comportements identifiés.	Énonce clairement 3 liens scientifiquement démontrés entre l'activité physique et la santé. Énonce clairement 3 comportements (dont au moins 1 positif et 1 négatif) affectant sa propre santé. Identifie clairement l'effet ressenti physiologiquement pour chacun des 3 comportements identifiés.	Énonce clairement 4 liens scientifiquement démontrés entre l'activité physique et la santé. Énonce clairement 4 comportements (dont au moins 1 positif et 1 négatif) affectant sa propre santé. Identifie clairement l'effet ressenti physiologiquement pour chacun des 4 comportements identifiés.
Critère 2 - Présentation des évaluations de son niveau de condition physique			
1	2 (seuil de réussite)	3	4
Présente partiellement les résultats des tests d'évaluation de la condition physique et utilise rarement les bonnes unités de mesure.	Présente les résultats de tous les tests d'évaluation de la condition physique et utilise la plupart du temps les bonnes unités de mesure.	Présente clairement les résultats de tous les tests d'évaluation de la condition physique sur les fiches originales de l'A.A. 4 et utilise généralement les bonnes unités de mesure. A.A. = activité d'apprentissage	Présente clairement les résultats de tous les tests d'évaluation de la condition physique sur les fiches originales de l'A.A. 4 et utilise toujours les bonnes unités de mesure. La présentation est particulièrement soignée.

Critère 3 - Identification des besoins et motivations			
1	2 (seuil de réussite)	3	4
<p>Identifie ses besoins partiellement en fonction de ses évaluations et peu en lien avec les données scientifiques qui précisent la condition physique souhaitable.</p> <p>Les justifications des besoins ne sont pas cohérentes et précises en fonction des comportements positifs et négatifs rapportés.</p> <p>Présente de façon artificielle ses motivations personnelles en vue d'améliorer sa pratique d'activité physique favorisant la santé.</p> <p>Il manque nettement de cohérence dans la présentation de ses besoins et de ses motivations.</p>	<p>Identifie ses besoins en fonction de ses évaluations et généralement en lien avec les données scientifiques qui précisent la condition physique souhaitable.</p> <p>Les justifications des besoins ne sont pas toujours cohérentes et précises en fonction des comportements positifs et négatifs rapportés.</p> <p>Présente de façon peu convaincante ses motivations personnelles en vue d'améliorer sa pratique d'activité physique favorisant la santé.</p>	<p>Identifie clairement ses besoins en fonction de ses évaluations et en lien avec les données scientifiques qui précisent la condition physique souhaitable.</p> <p>Les justifications des besoins sont cohérentes et précises en fonction des comportements positifs et négatifs rapportés.</p> <p>Présente de façon convaincante ses motivations personnelles en vue d'améliorer son sa pratique d'activité physique favorisant la santé.</p>	<p>Identifie très clairement ses besoins en fonction de ses évaluations et en lien avec les données scientifiques qui précisent la condition physique souhaitable.</p> <p>Les justifications des besoins sont cohérentes et précises en fonction des comportements positifs et négatifs rapportés. L'ensemble des liens est particulièrement bien exprimé.</p> <p>Présente de façon très convaincante ses motivations personnelles en vue d'améliorer sa condition physique, de plus, l'évalué montre sa volonté de réussir à améliorer ou maintenir, s'il y a lieu) sa pratique d'activité physique favorisant la santé.</p>
Critère 4 - Présentation du mode d'action;			
1	2 (seuil de réussite)	3	4
<p>Le plan d'action n'est pas présenté clairement et de façon détaillée.</p> <p>Il est très peu cohérent avec l'identification des besoins.</p> <p>Il est très peu réaliste en fonction du temps et du niveau de sa condition physique de départ.</p> <p>La réussite du plan d'action sera très difficilement observable et mesurable (trop théorique).</p> <p>L'ensemble du plan d'action est négligé.</p>	<p>Le plan d'action n'est pas toujours présenté clairement et de façon détaillée.</p> <p>Il est partiellement cohérent avec l'identification des besoins.</p> <p>Il est peu réaliste en fonction du temps et du niveau de sa condition physique de départ.</p> <p>La réussite du plan d'action sera difficilement observable et mesurable (trop théorique).</p>	<p>Le plan d'action est présenté clairement et de façon détaillée.</p> <p>Il est très cohérent avec l'identification des besoins.</p> <p>Il est réaliste en fonction du temps et du niveau de sa condition physique de départ.</p> <p>La réussite du plan d'action sera observable et mesurable (pas seulement théorique).</p> <p>Utilise généralement la terminologie usuelle en éducation physique.</p>	<p>Le plan d'action est présenté clairement et de façon détaillée. La présentation est particulièrement minutieuse.</p> <p>Il est très cohérent avec l'identification des besoins.</p> <p>Il est réaliste en fonction du temps et du niveau de sa condition physique de départ. Il identifie les contraintes possibles à sa réussite. Il fait un lien avec ses motivations personnelles.</p> <p>La réussite du plan d'action sera facilement observable et mesurable (pas seulement théorique).</p> <p>Respecte rigoureusement la terminologie usuelle en éducation physique.</p>

Critère 5 - Présentation de la mise en œuvre du plan d'action			
1	2 (seuil de réussite)	3	4
<p>Présentation non acceptable des fiches pour le compte rendu de la pratique hebdomadaire de l'activité physique réalisée.</p> <p>Le compte rendu ne distingue pas le type d'activité en rapport avec la qualité physique recherchée.</p> <p>La nature et la quantité des activités réalisées sont peu réalistes par rapport au contexte et à l'intensité de son entraînement.</p> <p>Les modes d'entraînement reconnus sont peu respectés pour chaque qualité physique.</p> <p>On n'observe pas de progression dans l'intensité des activités.</p> <p>S'il y a eu réajustement du plan d'action, les justifications et les explications sont inadéquates.</p> <p>Utilise rarement la terminologie usuelle en éducation physique.</p> <p>En classe, il fait preuve de très peu d'engagement et d'une participation très peu active.</p>	<p>Présentation acceptable des fiches pour le compte rendu de la pratique hebdomadaire de l'activité physique réalisée.</p> <p>Le compte rendu distingue le type d'activité en rapport avec la qualité physique recherchée.</p> <p>La nature et la quantité des activités réalisées sont assez réalistes par rapport au contexte.</p> <p>Les modes d'entraînement reconnus sont généralement respectés pour chaque qualité physique.</p> <p>On peut observer une légère progression dans l'intensité des activités.</p> <p>S'il y a eu réajustement du plan d'action, les justifications et les explications sont plus ou moins justes.</p> <p>N'utilise pas suffisamment la terminologie usuelle en éducation physique.</p> <p>En classe, il fait preuve de peu d'engagement et d'une participation peu active.</p>	<p>Présentation claire des fiches pour le compte rendu de la pratique hebdomadaire de l'activité physique réalisée.</p> <p>Le compte rendu distingue clairement le type d'activité en rapport avec la qualité physique recherchée.</p> <p>La nature et la quantité des activités réalisées sont réalistes par rapport au contexte.</p> <p>Les modes d'entraînement reconnus sont respectés pour chaque qualité physique.</p> <p>On peut observer une progression dans l'intensité des activités.</p> <p>S'il y a eu réajustement du plan d'action, les justifications et les explications sont justes.</p> <p>Utilise généralement la terminologie usuelle en éducation physique.</p> <p>En classe, il fait preuve d'un engagement sérieux et d'une participation active.</p>	<p>Présentation claire et précise des fiches pour le compte rendu de la pratique hebdomadaire de l'activité physique réalisée. Le compte rendu distingue clairement le type d'activité en rapport avec la qualité physique recherchée.</p> <p>La nature et la quantité des activités réalisées sont réalistes par rapport au contexte.</p> <p>Les modes d'entraînement reconnus sont respectés pour chaque qualité physique.</p> <p>On peut observer une forte progression dans l'intensité des activités.</p> <p>S'il y a eu réajustement du plan d'action, les justifications et les explications sont justes et très pertinentes.</p> <p>Respecte rigoureusement la terminologie usuelle en éducation physique.</p> <p>En classe, il fait preuve d'un engagement remarquable et d'une participation très active. Démonstre une forte curiosité ou un haut degré d'autonomie.</p>

Critère 6 - Application des règles de sécurité			
1	2 (seuil de réussite)	3	4
<p>Manque régulièrement aux règles de sécurité dans sa pratique de l'activité physique.</p> <p>Respecte peu ses limites personnelles</p> <p>Respecte peu les normes de l'activité ou des activités pratiquées.</p> <p>Commets assez souvent des imprudences. Exécute régulièrement des exercices dangereux ou non recommandés. En général, montre peu d'intérêt pour les règles de sécurité.</p>	<p>Application assez constante des règles de sécurité dans sa pratique de l'activité physique.</p> <p>Respecte généralement ses limites personnelles.</p> <p>Respecte généralement les normes de l'activité ou des activités pratiquées.</p> <p>Commets à l'occasion des imprudences. Exécute quelques fois des exercices dangereux ou non recommandés.</p>	<p>Application constante des règles de sécurité dans sa pratique de l'activité physique.</p> <p>Respecte ses limites personnelles.</p> <p>Respecte les normes de l'activité ou des activités pratiquées.</p> <p>Commets rarement des imprudences. Exécute rarement des exercices dangereux ou non recommandés.</p>	<p>Application constante et rigoureuse de règles de sécurité dans sa pratique de l'activité physique. Porte une attention particulière à la sécurité par rapport au matériel utilisé.</p> <p>Respecte ses limites personnelles.</p> <p>Respecte rigoureusement les normes de l'activité ou des activités pratiquées.</p> <p>À toujours un comportement sécuritaire. N'exécute jamais d'exercices dangereux ou non-recommandés</p>
Critère 7 - Analyse critique du plan d'action et de sa réalisation			
1	2 (seuil de réussite)	3	4
<p>Présente très partiellement une évaluation de sa démarche en précisant peu les réussites et les difficultés éprouvées, et identifie mal les causes de difficultés et les facteurs de réussite vécus.</p> <p>L'analyse est loin d'être juste et précise, car il manque beaucoup d'éléments</p> <p>Montre très peu d'objectivité dans son analyse.</p> <p>L'analyse n'est nettement pas rigoureuse et est très peu détaillée.</p> <p>Identifie peu ou pas les effets positifs et négatifs ressentis à la suite de la réalisation de son plan d'action.</p> <p>L'ensemble de l'analyse critique est négligé.</p>	<p>Présente partiellement une évaluation de sa démarche en précisant les réussites et les difficultés éprouvées, ainsi que les causes de difficultés et les facteurs de réussite vécus.</p> <p>L'analyse n'est pas toujours juste et précise.</p> <p>Montre peu d'objectivité dans son analyse.</p> <p>L'analyse est peu rigoureuse et peu détaillée.</p> <p>Identifie peu les effets positifs et négatifs ressentis à la suite de la réalisation de son plan d'action.</p>	<p>Présente une évaluation de sa démarche en précisant les réussites et les difficultés éprouvées, ainsi que les causes de difficultés et les facteurs de réussite vécus.</p> <p>Identifie des moyens qui auraient pu rendre sa démarche plus efficace, s'il y a lieu.</p> <p>Montre une bonne objectivité dans son analyse.</p> <p>L'analyse est rigoureuse et détaillée.</p> <p>Identifie les effets positifs et négatifs ressentis à la suite de la réalisation de son plan d'action.</p>	<p>Présente une évaluation de sa démarche en précisant les réussites et les difficultés éprouvées, ainsi que les causes de difficultés et les facteurs de réussite vécus.</p> <p>Identifie des moyens qui auraient pu rendre sa démarche plus efficace, s'il y a lieu.</p> <p>Montre une grande objectivité dans son analyse et fait preuve de créativité.</p> <p>L'analyse est très rigoureuse et très détaillée.</p> <p>Identifie très clairement les effets positifs et négatifs ressentis à la suite de la réalisation de son plan d'action.</p> <p>Dans l'ensemble, son analyse critique montre un grand sens critique.</p>

ÉVALUATION GLOBALE DE LA COMPÉTENCE

1	2	3	4
<i>Acquisition insuffisante</i>	<i>Acquisition minimale (seuil)</i>	<i>Acquisition attendue</i>	<i>Acquisition au-delà de l'attente</i>
0 À 59	60 À 69	70 À 89	90 À 100
- + _____	- + _____	- + _____	- + _____

ÉVALUATION GLOBALE DE LA COMPÉTENCE

NOTE : _____/100

Consignes pour l'évaluateur :

- 1) Encerclez la lettre correspondant à l'échelon qui situe le mieux les actions de l'élève en fonction de la description de chaque échelon. Puis, notez chaque critère en fonction des intervalles proposés. (*Vous pouvez aussi pondérer les critères, s'il y a lieu*)
- 2) À partir de l'évaluation de chaque critère, déterminez une note globale pour la compétence en prenant en compte l'ensemble de critères, sans les additionner. L'évaluation globale de la compétence (ou du cours) ne représente pas la moyenne ou la sommation des résultats de chaque critère, bien que cette évaluation soit fortement influencée par ces résultats. Si l'évaluateur est inconfortable avec la façon précédente de noter la compétence, il peut se rassurer en additionnant les critères d'évaluation de la compétence, après les avoir pondérés pour obtenir une note sur 100 points.
- 3) L'utilisation de telle grille prend du sens seulement si nous avons utilisé des grilles d'évaluation formative préalablement. L'étudiant doit bien comprendre la grille utilisée. La grille est donnée à l'étudiant le plus tôt dans la session (idéalement, dans le plan de cours).
- 4) L'évalué doit être informé clairement du processus d'évaluation. Étant donné la nouveauté de ce processus, il est nécessaire d'expliquer et de discuter de cette forme d'évaluation avec les évalués, et ce, à plusieurs reprises.

Annexe 4.

Les questionnaires français et anglais

Comité des enseignantes et enseignants d'éducation physique au collégial

SONDAGE SUR LES PRATIQUES ÉVALUATIVES DES ENSEIGNANTS ET ENSEIGNANTES EN ÉDUCATION PHYSIQUE AU COLLÉGIAL

Aux enseignants et enseignantes en éducation physique du réseau collégial,

À la suite des nombreux questionnements soulevés au cours des dernières années concernant l'évaluation des apprentissages, le *comité des enseignantes et enseignants d'éducation physique au collégial* en collaboration avec la FÉÉPEQ, nous a confié le mandat de réaliser un sondage qui permettra de dresser un inventaire de toutes les modalités et de tous les moyens d'évaluation qui sont susceptibles d'être utilisés dans le cadre de votre travail d'enseignant et d'enseignante. Le fait de répondre au questionnaire sera pour certains une occasion de réfléchir à leurs pratiques d'évaluation, pour d'autres, de se rendre compte de l'ensemble des possibilités qui peuvent exister en évaluation dans le réseau collégial. Les résultats de ce sondage nous permettront de mieux identifier les difficultés rencontrées en évaluation des apprentissages et à identifier des stratégies qui pourraient aider les enseignants et enseignantes en éducation physique dans leurs tâches d'évaluation.

Comme les modalités et les moyens d'évaluation utilisés risquent d'être fort variés d'un enseignant à l'autre, nous avons dû élaborer un instrument très exhaustif et qui, à première vue, pourra vous sembler un peu long. Étant donné l'importance de ce sondage pour vous et l'ensemble de notre communauté, nous sommes assurés que vous trouverez la motivation pour répondre en grand nombre à ce questionnaire (durée : environ 30 minutes).

Les réponses à ce questionnaire seront traitées de façon anonyme et confidentielle. L'équipe de recherche s'engage à respecter les règles déontologiques en vigueur pour une recherche scientifique.

Nous vous demandons de retourner votre questionnaire au plus tard le **8 décembre 2011**.

Les résultats de ce sondage seront présentés lors de la prochaine rencontre du *comité des enseignantes et enseignants d'éducation physique au collégial* (juin 2012). Une synthèse des résultats, ainsi qu'une interprétation du sondage seront expédiées à tous les enseignants et enseignantes en éducation physique au niveau collégial.

Votre participation est importante, merci pour votre précieuse collaboration,

Le comité responsable du sondage :

Gilles Champagne, Luc Chiasson, Robert Delaney, François Lasnier

POUR ACCÉDER AU QUESTIONNAIRE CLIQUEZ SUR LE LIEN SUIVANT :

Inventaire des pratiques évaluatives des enseignants et enseignantes en éducation physique au collégial

CONSIGNES

Avant de répondre au questionnaire, nous vous demandons de **choisir** une des trois compétences du programme d'éducation physique. Choisissez la compétence qui représente la **majorité de vos cours**. Une fois ce choix fait, vous devez répondre aux questions en référant **uniquement** à la compétence choisie. Ainsi, il vous sera plus facile de répondre à l'ensemble des questions.

Sélectionnez la compétence (ou l'ensemble) qui vous servira de référence pour toutes vos réponses.

- Ensemble 1 (4EP0) *Analyser sa pratique de l'activité physique au regard des habitudes de vie favorisant la santé.*
- Ensemble 2 (4EP1) *Améliorer son efficacité lors de la pratique d'une activité physique.*
- Ensemble 3 (4EP2) *Démontrer sa capacité à prendre en charge sa pratique de l'activité physique dans une perspective de santé.*

Pour coter chaque item, utilisez l'échelle de mesure suivante :

Échelle de fréquence de type Likert à 6 cotes

- 1 = Jamais,
- 2 = Très rarement
- 3 = Rarement
- 4 = Assez souvent
- 5 = Souvent
- 6 = Toujours

→Indiquez le chiffre qui correspond à votre choix.

Les concepts retenus pour composer l'inventaire des pratiques évaluatives des enseignants et des enseignantes du collégial sont les suivants :

- A. MODALITÉS DES ÉVALUATIONS : façons particulières d'évaluer
- B. INSTRUMENTS DE MESURE : outils fabriqués (examen théorique, travail long, recherche, portfolio, carnet de bord, tests physiques ou autres) qui permettent de mesurer les apprentissages et les productions de l'étudiant. INSTRUMENTS DE JUGEMENT : outils fabriqués permettant de noter ou de coter une production, tels une liste des critères d'évaluation, un barème de notation, des grilles de correction, des grilles d'évaluation descriptives.
- C. COMPOSITION DE LA NOTE : manière de déterminer le résultat final d'un étudiant en assemblant plusieurs éléments d'évaluation
- D. QUALITÉS DES ÉVALUATIONS : aspect sensible et non mesurable des évaluations.

Dans le questionnaire, afin d'en alléger la présentation, le masculin réfère aux deux genres.

Dans ce questionnaire, tous les items se rapportent à l'**évaluation sommative**, et non à l'évaluation formative. Étant donné la longueur du présent questionnaire, nous avons décidé de reporter à une date ultérieure la section sur l'évaluation formative.

Il n'y a pas de bonnes ou de mauvaises réponses. Toutefois, il est important de répondre à toutes les questions.

A. MODALITÉS DES ÉVALUATIONS

Ce concept regroupe les objets (sujets) d'évaluation et les façons particulières d'évaluer ces objets d'évaluation.

ITEMS	FRÉQUENCE						
	Jamais	1	2	3	4	5	6
1. J'évalue les connaissances disciplinaires (savoirs, connaissances déclaratives).	1	2	3	4	5	6	
2. J'évalue les attitudes (ex. : détermination, engagement).	1	2	3	4	5	6	
3. J'évalue les comportements (ex. : ponctualité).	1	2	3	4	5	6	
4. J'évalue les habiletés motrices.	1	2	3	4	5	6	
5. J'évalue les habiletés intellectuelles.	1	2	3	4	5	6	
6. J'évalue les performances physiques (ex. : nombre de lancers réussis, technique de la manchette au volley-ball).	1	2	3	4	5	6	
7. J'évalue les performances physiques en situation de jeu (ex. : application des stratégies collectives en jeu).	1	2	3	4	5	6	
8. J'évalue la condition physique avec des tests standardisés (ex. : .physitest canadien, redressements assis).	1	2	3	4	5	6	
9. J'évalue la condition physique sans tests standardisés (test sans normes de références).	1	2	3	4	5	6	
10. J'évalue les éléments de la compétence en utilisant les critères de performance tels qu'écrits dans le devis ministériel.	1	2	3	4	5	6	
11. Je réfère aux éléments de la compétence pour rédiger mes propres critères d'évaluation (différents des critères de performance)	1	2	3	4	5	6	
12. J'évalue tous les éléments de la compétence.	1	2	3	4	5	6	
13. Mes évaluations permettent de juger du degré de développement de la compétence.	1	2	3	4	5	6	
14. Je fais des évaluations qui font appel au par cœur.	1	2	3	4	5	6	
15. Lors de mes évaluations, les étudiants ont accès à leurs notes de cours et aux autres documents utilisés dans le cours.	1	2	3	4	5	6	
16. Pour un même cours, j'utilise les mêmes critères (quelques soient les critères) que ceux de mes collègues.	1	2	3	4	5	6	
17. Pour un même cours, j'utilise les mêmes instruments d'évaluation que ceux de mes collègues.	1	2	3	4	5	6	
18. Pour un même cours, j'applique les mêmes conditions d'évaluation, que celles de mes collègues.	1	2	3	4	5	6	
19. J'informe mes étudiants, à l'avance, sur quoi portera l'évaluation.	1	2	3	4	5	6	
20. J'informe mes étudiants, à l'avance, des critères d'évaluation.	1	2	3	4	5	6	
21. J'informe mes étudiants, à l'avance, de la pondération des critères d'évaluation.	1	2	3	4	5	6	
22. Je me conforme aux directives du département lors de mes évaluations.	1	2	3	4	5	6	
23. J'utilise des moyens de détection des cas de plagiat.	1	2	3	4	5	6	
24. Je mets en place des moyens pour prévenir le plagiat.	1	2	3	4	5	6	

25. Je sanctionne tout plagiat.	1	2	3	4	5	6
26. Je fais 3 évaluations et moins par session.	1	2	3	4	5	6
27. Je fais 4 évaluations et plus par session .	1	2	3	4	5	6
28. Je répartir également les points de mon évaluation sur l'ensemble de la session (ex. : 50 % des points avant la mi-session et 50 % après la mi-session).	1	2	3	4	5	6
29. Les points de mes évaluations sont répartis également sur mes critères d'évaluation.	1	2	3	4	5	6
30. Je pondère mes critères d'évaluation selon leur importance.	1	2	3	4	5	6
31. Pour mon évaluation, j'applique la règle du double seuil suivante : 60 % sur l'ensemble des évaluations (incluant l'épreuve finale) et en plus avoir au moins 60 % à l'épreuve finale.	1	2	3	4	5	6
32. Pour mon évaluation, j'applique la règle du double seuil suivante : avoir déjà obtenu 60 % pour accéder à l'évaluation finale et obtenir au moins 60 % à l'évaluation finale.	1	2	3	4	5	6
33. Pour mon évaluation, si un étudiant n'obtient pas 60 % pour l'ensemble de ses évaluations, mais qu'il obtient 60 % et plus à l'épreuve finale, il réussit automatiquement le cours.	1	2	3	4	5	6
34. Je fais certaines évaluations qui comptent pour moins de 12 % des points de la session.	1	2	3	4	5	6
35. Mon épreuve épreuve-synthèse de cours (épreuve finale) compte pour au moins 40 % des points de la session.	1	2	3	4	5	6

B. INSTRUMENTS DE MESURE

Dans ce questionnaire, le terme « instrument de mesure » réfère aux différents instruments qui servent à recueillir les productions de l'étudiant, telles que des examens, des travaux, des recherches, un portfolio, un carnet de bord, des tests physiques ou autres. Les instruments de mesure ne conduisent pas directement à une note; ce sont les « instruments de jugement » qui permettent d'attribuer une note ou une cote à une production.

ITEMS	FRÉQUENCE					
	Jamais				Toujours	
1. J'utilise des tests de condition physique dans mes évaluations.	1	2	3	4	5	6
2. J'utilise des tests d'efficacité technique (ou d'habileté motrice) dans mes évaluations.	1	2	3	4	5	6
3. Pour mon évaluation, mes étudiants produisent un portfolio qui fait la synthèse de leurs travaux pendant la session.	1	2	3	4	5	6
4. Pour mon évaluation, mes élèves produisent un carnet de bord pour décrire leur travail quotidien ou hebdomadaire.	1	2	3	4	5	6
5. Pour mon évaluation, j'utilise des fiches d'observation.	1	2	3	4	5	6
6. Pour mon évaluation, je demande aux étudiants de réaliser un ou des projets.	1	2	3	4	5	6
7. J'utilise des travaux longs pour mon évaluation (une page et plus).	1	2	3	4	5	6
8. Pour mon évaluation, j'utilise des études de cas.	1	2	3	4	5	6
9. Pour mon évaluation, j'utilise des travaux d'équipe.	1	2	3	4	5	6
10. Pour mon évaluation, j'utilise des exposés oraux.	1	2	3	4	5	6

11. J'utilise un instrument de mesure pour évaluer les attitudes.	1	2	3	4	5	6
12. Pour mon évaluation, je fais des entrevues.	1	2	3	4	5	6
13. Pour mon évaluation, je fais des examens théoriques écrits.	1	2	3	4	5	6
14. Dans mes examens théoriques, j'utilise des questions à réponse longue (ou à développement).	1	2	3	4	5	6
15. Dans mes examens théoriques, j'utilise des questions à réponse courte.	1	2	3	4	5	6
16. Dans mes examens théoriques, j'utilise des questions à choix multiple.	1	2	3	4	5	6
17. Dans mes examens théoriques, j'utilise des questions à réponse dichotomique (vrai ou faux).	1	2	3	4	5	6
18. Dans mes examens théoriques, j'utilise des questions au choix (ex. : répondre à 2 questions sur 4 proposées).	1	2	3	4	5	6
19. Dans mes examens théoriques, j'utilise des questions d'appariement (associer des termes à d'autres termes).	1	2	3	4	5	6
20. Lorsque j'utilise un instrument de mesure théorique ou pratique, je m'assure qu'il est en lien direct avec la compétence.	1	2	3	4	5	6
21. Lorsque j'utilise un instrument de mesure, j'explique aux étudiants son lien avec la compétence.	1	2	3	4	5	6
22. J'évalue la compétence avec des tâches d'intégration (tâches sollicitant tous les éléments de la compétence).	1	2	3	4	5	6
23. Mes tâches d'évaluation sollicitent seulement une partie des éléments de la compétence.	1	2	3	4	5	6
24. Je fais une épreuve-synthèse de cours (épreuve finale).	1	2	3	4	5	6

C. INSTRUMENTS DE JUGEMENT

Dans ce questionnaire, le terme « instrument de jugement » réfère aux outils permettant de noter ou de coter une production, tels une liste des critères d'évaluation, un barème de notation, des grilles de correction, des grilles d'évaluation descriptives.

ITEMS	FRÉQUENCE					
	Jamais				Toujours	
1. Pour porter un jugement sur le degré de réussite de l'étudiant, je fais une sélection parmi les critères de performance de la compétence du devis.	1	2	3	4	5	6
2. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise tous les critères de performance de la compétence du devis.	1	2	3	4	5	6
3. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise des critères autres que les critères de performance de la compétence du devis.	1	2	3	4	5	6
4. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise les éléments de la compétence du devis comme critères d'évaluation de la compétence.	1	2	3	4	5	6
5. Pour porter un jugement sur le degré de réussite de l'étudiant, je rédige moi-même les critères d'évaluation de la compétence.	1	2	3	4	5	6

6. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise des critères d'évaluation sélectionnés par le département.	1	2	3	4	5	6
7. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise une grille d'évaluation décrivant les résultats attendus pour chaque critère d'évaluation et chaque niveau de réussite.	1	2	3	4	5	6
8. Pour porter un jugement sur le degré de réussite de l'étudiant, j'utilise une grille d'évaluation décrivant les résultats attendus pour chaque niveau de réussite, en considérant la compétence globalement (c.-à-d. sans spécifier chaque critère).	1	2	3	4	5	6
9. Je note à partir d'une cote (ex. : A, B, C, D, E).	1	2	3	4	5	6
10. Pour porter un jugement sur les attitudes (ex. : détermination, engagement) de l'étudiant, j'utilise une grille d'évaluation décrivant les résultats attendus.	1	2	3	4	5	6

D. COMPOSITION DE LA NOTE

Ce concept regroupe les différentes façons de composer les notes et la répartition des points à attribuer.

ITEMS	FRÉQUENCE					
	Jamais					Toujours
1. J'accorde des points pour les résultats individuels dans des tournois compétitifs.	1	2	3	4	5	6
2. J'accorde des points pour les résultats d'équipe dans des tournois compétitifs (ex. : 2 points pour une victoire).	1	2	3	4	5	6
3. Je donne des points pour des réalisations collectives (ex. : travaux d'équipe, chorégraphies).	1	2	3	4	5	6
4. J'attribue des points à chaque critère de performance de la compétence.	1	2	3	4	5	6
5. J'attribue des points à chaque élément de la compétence.	1	2	3	4	5	6
6. J'utilise les critères de performance de la compétence pour noter l'épreuve finale.	1	2	3	4	5	6
7. J'utilise les éléments de la compétence pour noter l'épreuve finale.	1	2	3	4	5	6
8. J'utilise mes propres critères d'évaluation pour noter l'épreuve finale.	1	2	3	4	5	6
9. J'accorde des points selon les différents degrés de maîtrise de la compétence (échelons), établis dans ma grille descriptive.	1	2	3	4	5	6
10. J'accorde des points pour les présences aux cours.	1	2	3	4	5	6
11. J'enlève des points pour les absences au cours.	1	2	3	4	5	6
12. Trop d'absences entraînent automatiquement un échec.	1	2	3	4	5	6
13. J'accorde des points pour la participation à des activités hors cours.	1	2	3	4	5	6
14. Une partie des points de l'évaluation provient de l'évaluation par les pairs.	1	2	3	4	5	6
15. Une partie des points de l'évaluation provient de l'évaluation par les pairs, combinée à la mienne.	1	2	3	4	5	6
16. Une partie des points de l'évaluation provient d'une auto-évaluation.	1	2	3	4	5	6

17. Une partie des points de l'évaluation provient d'une coévaluation (enseignant-étudiant).	1	2	3	4	5	6
18. J'accorde des points pour l'effort.	1	2	3	4	5	6
19. Je suis le seul à donner des notes à mes étudiants.	1	2	3	4	5	6
20. Je donne des points pour le progrès, sans utiliser un prétest et un post-test.	1	2	3	4	5	6
21. Je donne des points pour le progrès, en utilisant un prétest et un post-test.	1	2	3	4	5	6
22. J'accorde des points pour les attitudes (ex. : détermination, engagement).	1	2	3	4	5	6
23. J'accorde des points pour les comportements (ex. : ponctualité).	1	2	3	4	5	6
24. J'enlève des points pour les comportements inappropriés.	1	2	3	4	5	6
25. Je donne des points pour la qualité de la langue écrite.	1	2	3	4	5	6
26. J'enlève des points pour un usage inapproprié de la langue écrite.	1	2	3	4	5	6
27. Je donne spécifiquement des points pour le critère de performance « <i>communication claire et argumentée de sa proposition d'activités physique</i> ».	1	2	3	4	5	6

E. QUALITÉS DES ÉVALUATIONS

Ce concept énumère une série de pratiques qui visent à assurer certaines qualités lorsqu'on évalue des étudiants.

ITEMS	FRÉQUENCE					
	Jamais				Toujours	
1. Mes instruments évaluent précisément les objets d'évaluation prévus dans le plan de cours.	1	2	3	4	5	6
2. La nature de mes évaluations est décrite dans le plan de cours (ex. : travaux, examens, évaluations pratiques).	1	2	3	4	5	6
3. Mes modalités d'évaluation sont décrites dans le plan de cours (ex. : critères d'évaluation, pondération des critères, grille d'évaluation, conditions de passation).	1	2	3	4	5	6
4. Mes consignes pour l'évaluation sont faciles à comprendre par les étudiants.	1	2	3	4	5	6
5. Mes consignes d'évaluation sont écrites.	1	2	3	4	5	6
6. Mes consignes d'évaluation sont données seulement verbalement.	1	2	3	4	5	6
7. Mes consignes d'évaluation sont connues de l'étudiant avant l'évaluation.	1	2	3	4	5	6
8. Mes évaluations permettent de discriminer les étudiants forts des étudiants faibles.	1	2	3	4	5	6
9. Mes évaluations visent à comparer les étudiants entre eux (évaluations normatives).	1	2	3	4	5	6
10. Mes évaluations visent à comparer les résultats des étudiants à des résultats attendus.	1	2	3	4	5	6
11. J'ai un souci d'équité dans mes évaluations.	1	2	3	4	5	6

12. La force de la classe influence mes évaluations.	1	2	3	4	5	6
13. Les résultats passés d'un étudiant influencent ses futures évaluations.	1	2	3	4	5	6
14. J'accorde une note de passage à un étudiant qui est près de la note de passage.	1	2	3	4	5	6
15. J'hésite à faire échouer un étudiant.	1	2	3	4	5	6
16. Le jugement de mes collègues influence le niveau de sévérité de mes évaluations.	1	2	3	4	5	6
17. La distribution de mes notes suit la courbe normale (forme d'une cloche).	1	2	3	4	5	6
18. Je fais valider mes instruments de mesure par des collègues.	1	2	3	4	5	6
19. Je fais valider mes instruments de jugement par des collègues.	1	2	3	4	5	6
20. Les comportements d'un étudiant influencent son évaluation.	1	2	3	4	5	6
21. Les tâches d'intégration (tâches sollicitant tous les éléments de la compétence) qui servent à l'évaluation sont inscrites dans mon plan de cours.	1	2	3	4	5	6
22. Je m'assure que mes tâches d'intégration sont significatives pour les étudiants.	1	2	3	4	5	6
23. Mes évaluations permettent de discriminer les étudiants qui maîtrisent bien la compétence par rapport aux autres.	1	2	3	4	5	6
24. Je m'assure que mes grilles d'évaluation ou autres instruments de jugement sont faciles à comprendre.	1	2	3	4	5	6

AUTRES INFORMATIONS

CONSIGNES : Cochez les informations demandées, ou répondez dans l'espace prévu.

Sexe : 1 Homme 2 Femme

Nombre d'années d'enseignement au collégial : _____ ans

Formation : 1 baccalauréat

2 maîtrise 3 doctorat 4 autre _____ (Kinésiologie)

Depuis que vous enseignez au collégial combien d'heures de formation avez-vous suivi en évaluation des apprentissages ?

1 0 h 2 1h à 15 h 3 16 à 30 h 4 plus de 30 h

Évaluation sommative

Dans un même groupe, combien d'évaluations sommatives faites-vous en moyenne par session ?

À quelle fréquence faites-vous des évaluations formatives (formelles et informelles) ?

1 jamais 2 très rarement 3 rarement 4 souvent 5 assez souvent 6 très souvent

Pondération de la note

Combien de points sur 100 accordez-vous à l'évaluation finale (épreuve synthèse de cours) ?
_____/100 pts

Combien de points sur 100 accordez-vous à la théorie ? _____/100 pts

Sur l'ensemble de la session, combien de points sur 100 accordez-vous à la pratique ? _____/100 pts

Merci pour votre précieuse collaboration

Le comité responsable du sondage

Gilles Champagne, Luc Chiasson, Bob Delaney, François Lasnier

CEGEP Provincial Coordinating Committee

A SURVEY OF THE EVALUATION PRACTICES OF CEGEP PHYSICAL EDUCATION TEACHERS

To All CEGEP Physical Education Teachers,

As a result of the significant number of questions raised over the past years regarding the evaluation of learning, the Provincial Coordinating Committee, in conjunction with FÉÉPEQ, has mandated our committee to administer a survey that will help us identify the evaluation practices of CEGEP Physical Education teachers. This survey will permit us to better identify difficulties experienced in evaluation.

Given that evaluation practices tend to vary from one teacher to another and from one college to another, it has been necessary to develop a questionnaire that is as exhaustive in nature as possible. At first glance, this questionnaire may appear to you as being extremely long. However, we believe that the estimated time to answer all items is about 30 minutes. As a possible source of motivation, we believe that answering the questionnaire will provide you with the opportunity to reflect upon your own evaluation practices and perhaps discover new approaches or alternatives in time to be included in your courses next semester.

Given the importance of this survey to you as a teacher, as well as to our profession, we ask all of you to take the time to answer the questionnaire. We assure you that all answers will be dealt with in a confidential and anonymous manner. The research team is committed to respect the ethical rules for scientific research.

The results of this survey will be presented at the next Physical Education Provincial Coordinating Committee meeting that will be held in June of 2012. A summary of the results, as well as an interpretation of the findings will be sent to all CEGEP Physical Education teachers.

Please return your questionnaire no later than December 8, 2011.

Your participation is critical to the success of this survey. Thank you for your cooperation,

The Survey Committee:

Gilles Champagne, Luc Chiasson, Robert Delaney, François Lasnier

To access the questionnaire, please click on the following link :

An Inventory of the Evaluation Practices of CEGEP Physical Education Teachers

Instructions

In answering the questionnaire, please choose one of the three competencies (sets of objectives) from which the majority of your courses are taught. Once this choice is made, answer each statement referring only to this competency.

Identify below the competency that you will use in answering each statement.

- 4EP0 *To analyse physical activity from the standpoint of lifestyle behaviours that promote health*
- 4EP1 *To improve effectiveness when practicing a physical activity*
- 4EP2 *To demonstrate the ability to manage health enhancing physical activity*

In rating each item, use the following scale:

Frequency Scale with 6 Likert-type ratings

- 1 = Never
- 2 = Very Rarely
- 3 = Rarely
- 4 = Occasionally
- 5 = Frequently
- 6 = Always

The concepts selected to identify the evaluation practices of CEGEP Physical Education teachers are the following:

- A. Evaluation Procedures
- B. Measurement Tools
- C. Judgement Tools
- D. Makeup of the Grade
- E. Evaluation Quality

All items in this questionnaire deal only with summative and not formative evaluation. Given the length of the present questionnaire, we have decided to deal with formative evaluation at a later date.

Please note that there are no right or wrong answers to these questions. Keeping this in mind, please answer all questions to the best of your ability.

A. EVALUATION PROCEDURES

In this questionnaire the concept « Evaluation Procedures » refers to the intended matter to be evaluated or the object of the evaluation and the specific ways in which these are evaluated.

ITEMS	FREQUENCY					
	Never				Always	
36. I evaluate theoretical course content or disciplinary knowledge (knowledge, declarative knowledge).	1	2	3	4	5	6
37. I evaluate attitudes (ex: determination, commitment).	1	2	3	4	5	6
38. I evaluate behaviours (ex: punctuality).	1	2	3	4	5	6
39. I evaluate motor skills.	1	2	3	4	5	6
40. I evaluate intellectual skills.	1	2	3	4	5	6
41. I evaluate physical performance (ex: number of successful free throws, bump pass technique in volleyball).	1	2	3	4	5	6
42. I evaluate physical performance in game situations (ex: use of team strategies in a game situation).	1	2	3	4	5	6
43. I evaluate physical fitness using standardised tests. (ex: Canadian Standardized Fitness Test, , Sit-Up Test).	1	2	3	4	5	6
44. I evaluate physical fitness with non-standardised tests (test without reference norms).	1	2	3	4	5	6
45. I evaluate the elements of the competency using the performance criteria as they appear in the Ministerial Specifications (Devis).	1	2	3	4	5	6
46. I refer to the elements of the competency to write my own evaluation criteria. (different from the performance criteria)	1	2	3	4	5	6
47. I evaluate all the elements of the competency.	1	2	3	4	5	6
48. My evaluations permit me to judge the degree to which the student has developed the competency.	1	2	3	4	5	6
49. I carry out evaluations that require memorization.	1	2	3	4	5	6
50. During my evaluations, students have access to course notes and other documents used in the course.	1	2	3	4	5	6
51. For a given course, I use the same criteria (whatever these criteria may be) as my colleagues.	1	2	3	4	5	6
52. For a given course, I use the same evaluation tools as my colleagues.	1	2	3	4	5	6
53. For a given course, I apply the same evaluation conditions as my colleagues.	1	2	3	4	5	6
54. I inform my students in advance regarding what will be evaluated.	1	2	3	4	5	6
55. I inform my students in advance of the evaluation criteria.	1	2	3	4	5	6
56. I inform my students in advance of the weighting of the evaluation criteria.	1	2	3	4	5	6
57. I respect the departmental guidelines when evaluating.	1	2	3	4	5	6
58. I use ways to detect plagiarism.	1	2	3	4	5	6

59. I institute ways to prevent plagiarism.	1	2	3	4	5	6
60. I penalize all acts of plagiarism.	1	2	3	4	5	6
61. I administer 3 or fewer evaluations per semester.	1	2	3	4	5	6
62. I administer 4 or more evaluations per semester.	1	2	3	4	5	6
63. I distribute the marks equally over the semester (ex. 50% before mid-term and 50% after mid-term)	1	2	3	4	5	6
64. I distribute the marks equally over all of the evaluation criteria.	1	2	3	4	5	6
65. I weight the evaluation criteria according to their importance.	1	2	3	4	5	6
66. In administering my evaluations, I apply the following double threshold rule: 60% over all evaluations (including the final evaluation), as well as achieving at least 60% on the final test.	1	2	3	4	5	6
67. In administering my evaluations, I apply the following double threshold rule: having already achieved 60% in order to qualify to write the final and to obtain at least 60 % on the final.	1	2	3	4	5	6
68. In administering my evaluations, if a student does not achieve 60% over all evaluations but achieves 60% or more on the final, the student automatically passes the course.	1	2	3	4	5	6
69. I administer evaluations that are worth less than 12% of the final grade.	1	2	3	4	5	6
70. My final exam is worth at least 40% of the final grade.	1	2	3	4	5	6

B. MEASUREMENT TOOLS

In this questionnaire, the term « Measurement Tool » refers to the different tools used to collect the work done by the student, such as exams, assignments, research papers, portfolios, diaries, physical tests or others. The measurement tools do not directly lead to a grade; rather it is a « Judgement Tool » that enables the teacher to attribute a grade or a rating.

ITEMS	FREQUENCY					
	Never				Always	
25. I use fitness tests as part of my evaluation process.	1	2	3	4	5	6
26. I use skill tests as part of my evaluation process.	1	2	3	4	5	6
27. My students produce a portfolio that is a synthesis of the work done during the semester as part of my evaluation process.	1	2	3	4	5	6
28. My students keep a diary describing their daily or weekly work as part of my evaluation process.	1	2	3	4	5	6
29. I use observation forms as part of my evaluation process.	1	2	3	4	5	6
30. I have students carry out one or more projects as part of my evaluation process.	1	2	3	4	5	6
31. I use assignments of one or more pages as part of my evaluation process.	1	2	3	4	5	6
32. I use case studies as part of my evaluation process.	1	2	3	4	5	6
33. I use team work as part of my evaluation process.	1	2	3	4	5	6
34. I use oral presentations as part of my evaluation process.	1	2	3	4	5	6

35. I use a measurement tool to evaluate attitudes.	1	2	3	4	5	6
36. I carry out interviews as part of my evaluation process.	1	2	3	4	5	6
37. I give written theory exams as part of my evaluation process.	1	2	3	4	5	6
38. I use essay-type questions in my theory exams.	1	2	3	4	5	6
39. I use short-answer questions in my theory exams.	1	2	3	4	5	6
40. I use multiple-choice questions in my theory exams.	1	2	3	4	5	6
41. I use true or false questions in my theory exams.	1	2	3	4	5	6
42. I provide students with a choice of questions in my theory exams. (ex: answer 2 of the 4 proposed questions).	1	2	3	4	5	6
43. I use matching questions (associate one term with another) in my theory exams.	1	2	3	4	5	6
44. When I use either a theoretical or practical measurement tool, I ensure that it is directly related to the competency.	1	2	3	4	5	6
45. When I use a measurement tool, I explain its relationship with the competency to the students.	1	2	3	4	5	6
46. My integrative tasks be they theoretical or practical, solicit all the elements of the competency.	1	2	3	4	5	6
47. My integrative tasks (by definition, those that solicit all the elements of the competency), be they theoretical or practical, solicit only some of the elements of the competency.	1	2	3	4	5	6
48. Within my course, I carry out a comprehensive assessment (final exam).	1	2	3	4	5	6

C. JUDGEMENT TOOLS

In this questionnaire, the term « Judgement Tool » refers to the tools that enable you to judge or assess a student's response or performance in order to assign a grade or rating, such as a list of evaluation criteria, a rating scale, a correction grid (rubric), descriptive evaluation grid (rubric).

ITEMS	FREQUENCY					
	Never				Always	
11. To assess the degree of success achieved by the student, I select from amongst the performance criteria of the competency as defined in the Ministerial Specifications (Devis)	1	2	3	4	5	6
12. To assess the degree of success achieved by the student, I use all the performance criteria as defined in the Ministerial Specifications (Devis).	1	2	3	4	5	6
13. To assess the degree of success achieved by the student, I use criteria other than the performance criteria defined in the Ministerial Specifications (Devis).	1	2	3	4	5	6
14. To assess the degree of success achieved by the student, I use the elements of the competency as defined in the Ministerial Specifications (Devis) as evaluation criteria.	1	2	3	4	5	6
15. To assess the degree of success achieved by the student, I formulate evaluation criteria for the competency.	1	2	3	4	5	6

16. To assess the degree of success achieved by the student, I use the evaluation criteria selected by the Department.	1	2	3	4	5	6
17. To assess the degree of success achieved by the student, I use an evaluation grid (rubric) describing expected results for different levels of success for each evaluation criteria.	1	2	3	4	5	6
18. To assess the degree of success achieved by the student, I use an evaluation grid or rubric describing expected results for each level of success for the overall competency without specifying each element of the competency.	1	2	3	4	5	6
19. I grade using a rating scale (ex: A, B, C, D, E).	1	2	3	4	5	6
20. In order to assess student attitudes (ex: determination, commitment), I use an evaluation grid or rubric describing expected results.	1	2	3	4	5	6

D. MAKEUP OF THE GRADE

In this questionnaire, this concept refers to the different ways of composing grades and distributing marks.

ITEMS	FREQUENCY					
	Never				Always	
28. I give marks for individual results in a competitive tournament.	1	2	3	4	5	6
29. I give marks for team results in a competitive tournament (ex: 2 points for a victory)	1	2	3	4	5	6
30. I give marks for group achievements (ex: team work, choreographies).	1	2	3	4	5	6
31. I assign marks to each of the performance criteria of the competency.	1	2	3	4	5	6
32. I assign marks for each element of the competency.	1	2	3	4	5	6
33. I use the performance criteria of the competency to grade the final exam.	1	2	3	4	5	6
34. I use the elements of the competency to grade the final exam.	1	2	3	4	5	6
35. I use my own evaluation criteria to grade the final exam.	1	2	3	4	5	6
36. I give marks based upon the different degrees of attainment of the competency (scales), as established in a descriptive grid (rubric).	1	2	3	4	5	6
37. I give marks for attendance.	1	2	3	4	5	6
38. I deduct marks for absences.	1	2	3	4	5	6
39. Too many absences automatically result in a failure.	1	2	3	4	5	6
40. I give marks for participation in activities that take place outside of class.	1	2	3	4	5	6
41. Part of the grade comes from a peer evaluation.	1	2	3	4	5	6
42. Part of the grade comes from a peer evaluation combined with my own evaluation.	1	2	3	4	5	6
43. Part of the grade comes from a student self-evaluation.	1	2	3	4	5	6

44. Part of the grade comes from a co-evaluation (teacher-student).	1	2	3	4	5	6
45. I give marks for effort.	1	2	3	4	5	6
46. I am the sole contributor to the student grade.	1	2	3	4	5	6
47. I give marks for progress without the use of a pre- and post-test.	1	2	3	4	5	6
48. I give marks for progress using a pre- and post-test.	1	2	3	4	5	6
49. I give marks for attitude (ex: determination, commitment).	1	2	3	4	5	6
50. I give marks for behaviour (ex: punctuality).	1	2	3	4	5	6
51. I deduct marks for inappropriate behaviour.	1	2	3	4	5	6
52. I give marks for the quality of the written language.	1	2	3	4	5	6
53. I deduct marks for an inappropriate use of written language.	1	2	3	4	5	6
54. I assign marks for the performance criteria “ <i>Use of clear arguments to explain the choice of physical activity.</i> ”	1	2	3	4	5	6

E. EVALUATION QUALITY

In this questionnaire, this concept refers to a series of practices designed to ensure the quality of student evaluation.

ITEMS	FREQUENCY					
	Never				Always	
25. My tools precisely evaluate what they are intended to evaluate as set out in the course outline (syllabus).	1	2	3	4	5	6
26. The nature of my evaluation tools (ex: assignments, exams, practical evaluations) are described in the course outline (syllabus).	1	2	3	4	5	6
27. The evaluation procedures (ex: evaluation criteria, weighting of criteria, evaluation grids, requirements for passing) are described in the course outline (syllabus).	1	2	3	4	5	6
28. Students easily understand my evaluation instructions.	1	2	3	4	5	6
29. My evaluation instructions are written down.	1	2	3	4	5	6
30. My evaluation instructions are only given verbally.	1	2	3	4	5	6
31. Students know the evaluation instructions prior to the evaluation.	1	2	3	4	5	6
32. My evaluations differentiate (i.e.: discriminate) strong students from the weak ones.	1	2	3	4	5	6
33. My evaluations are designed to compare a student's performance to each other's. (normative evaluations).	1	2	3	4	5	6
34. My evaluations are designed to compare student results to expected results.	1	2	3	4	5	6
35. Evaluation fairness is a concern of mine.	1	2	3	4	5	6
36. The overall strength of the class influences my evaluations.	1	2	3	4	5	6

37. Past results of a student influence his/her future evaluations.	1	2	3	4	5	6
38. I grant a passing grade to a student who is close to a passing grade.	1	2	3	4	5	6
39. I hesitate to fail a student.	1	2	3	4	5	6
40. The judgement of my colleagues influences the level of difficulty of my evaluations.	1	2	3	4	5	6
41. My class grades follow a normal distribution curve.	1	2	3	4	5	6
42. My measurement tools are validated by my colleagues.	1	2	3	4	5	6
43. My judgement tools are validated by my colleagues.	1	2	3	4	5	6
44. A student's behaviour influences his/her evaluation.	1	2	3	4	5	6
45. Integration tasks (by definition, those that solicit all the elements of the competency) that are used for an evaluation are identified in my course outline (i.e: syllabus).	1	2	3	4	5	6
46. I ensure that my integration tasks are meaningful to the students.	1	2	3	4	5	6
47. My evaluations discriminate (i.e.: differentiate) students who have mastered the competency from the others.	1	2	3	4	5	6
48. I make sure that my evaluation grids or other judgement tools are easy to understand.	1	2	3	4	5	6

ADDITIONAL INFORMATION

Instructions: Please provide the following information:

Gender: 1 Male 2 Female

Indicate the number of years you have been teaching at the college level:

Indicate the numbers of years _____

Level of Education : 1 Undergraduate PE2 Masters PE 3 Doctorate PE
4 Undergraduate KIN 5 Masters KIN 6 Doctorate KIN 44
7 other (KIN)_____

Hours of Training

Since you began teaching at the college level, how many hours of training in evaluation of learning have you followed?

Indicate the number of hours

1 0 h 2 1-15 3 16-30 4 more than 30

SUMMATIVE EVALUATIONS

On average, how many summative evaluations do you carry out in each course during a semester?

Indicate the average number of summative evaluations per semester _____

Formative Evaluations

How often do you carry out either formal or informal formative evaluations?

1 never 2 very rarely 3 rarely 4 occasionally 5 frequently 6 always

WEIGHTING OF THE GRADE

How many marks out of 100 do you assign to the final exam? ____/100 marks

Over the semester, how many marks out of 100 are set aside for the theory? ____/100 marks

Over the semester, how many marks out of 100 are set aside for practical? ____/100 marks

Thank you for your invaluable assistance.

The Survey Committee

Gilles Champagne, Luc Chiasson, Robert Delaney, François Lasnier