

Le système québécois d'évaluation au niveau primaire : exploration des pénalisations possibles pour les étudiants des milieux défavorisés

Michel Carbonneau et Michel Trahan

Volume 6, numéro 2, printemps 1980

URI : <https://id.erudit.org/iderudit/900281ar>

DOI : <https://doi.org/10.7202/900281ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Revue des sciences de l'éducation

ISSN

0318-479X (imprimé)

1705-0065 (numérique)

[Découvrir la revue](#)

Citer cet article

Carbonneau, M. & Trahan, M. (1980). Le système québécois d'évaluation au niveau primaire : exploration des pénalisations possibles pour les étudiants des milieux défavorisés. *Revue des sciences de l'éducation*, 6(2), 231–248. <https://doi.org/10.7202/900281ar>

Résumé de l'article

Résultat d'une analyse théorique dont l'objectif était de découvrir si le système formel d'évaluation au niveau primaire risque de pénaliser les enfants des milieux défavorisés, le présent texte, délibérément partial, formule différentes hypothèses concernant divers aspects de l'évaluation éventuellement pénalisants pour lesdits milieux. Ainsi, les pénalités potentielles liées à l'utilisation de l'intelligence et de l'apprentissage en tant qu'objets d'évaluation sont discutées et illustrées par l'analyse de tests en usage dans le système scolaire québécois. De même, la priorité accordée à l'évaluation au service des décisions administratives par rapport à l'évaluation au service des apprentissages est établie clairement de même que les conséquences négatives de l'utilisation massive des résultats de l'évaluation pour des décisions administratives. Plus loin, l'approche normative est présentée comme la plus susceptible de pénaliser systématiquement les étudiants des milieux défavorisés alors que l'approche critériée semble moins contribuer à amplifier leur marginalité. Enfin, les implications d'une centralisation pédagogique sont analysées.

Le système québécois d'évaluation au niveau primaire : exploration des pénalisations possibles pour les étudiants des milieux défavorisés

Michel Carbonneau et Michel Trahan*

RÉSUMÉ

Résultat d'une analyse théorique dont l'objectif était de découvrir si le système formel d'évaluation au niveau primaire risque de pénaliser les enfants des milieux défavorisés, le présent texte, délibérément partial, formule différentes hypothèses concernant divers aspects de l'évaluation éventuellement pénalisants pour lesdits milieux. Ainsi, les pénalités potentielles liées à l'utilisation de l'intelligence et de l'apprentissage en tant qu'objets d'évaluation sont discutées et illustrées par l'analyse de tests en usage dans le système scolaire québécois. De même, la priorité accordée à l'évaluation au service des décisions administratives par rapport à l'évaluation au service des apprentissages est établie clairement de même que les conséquences négatives de l'utilisation massive des résultats de l'évaluation pour des décisions administratives. Plus loin, l'approche normative est présentée comme la plus susceptible de pénaliser systématiquement les étudiants des milieux défavorisés alors que l'approche critériée semble moins contribuer à amplifier leur marginalité. Enfin, les implications d'une centralisation pédagogique sont analysées.

* Carbonneau, Michel et Trahan, Michel : professeurs, Université de Montréal.

Introduction

Quelques considérations autour de la question de savoir s'il est possible, au niveau primaire, de procéder à des évaluations qui ne pénalisent pas les étudiants des milieux défavorisés constituent l'objet du présent texte. Nous tenterons donc d'identifier et d'analyser d'une façon plutôt théorique les éléments d'évaluation qui constituent des sources potentielles de désavantages pour les étudiants des milieux défavorisés. Pour ce faire, nous laisserons de côté certains éléments d'évaluation qui sont moins susceptibles de pénaliser ces étudiants pour nous concentrer sur les précédents et les décortiquer en amplifiant les sources de désavantages possibles. Une telle démarche vise surtout à élargir la problématique et à identifier des pistes de recherche plutôt qu'à présenter des résultats de recherches empiriques démontrant de tels effets pénalisants. Avant d'aborder plus directement la question, il nous paraît utile d'identifier deux grandes catégories de biais susceptibles de jouer à l'avantage ou au désavantage de certaines tranches de la population selon qu'ils originent soit des pratiques d'évaluation soit des contextes d'évaluation. Ainsi, certaines pratiques d'évaluation peuvent amplifier les différences entre les résultats des étudiants provenant des milieux défavorisés et ceux des autres : c'est le cas des tests normatifs qui tendent à maximiser les différences entre individus ou encore, celui des tests dont le contenu serait très fortement saturé de valeurs propres à un milieu en particulier. De même, les contextes d'évaluation — nous pensons surtout ici aux programmes d'études et aux attitudes et pratiques pédagogiques — peuvent contribuer à accentuer la marginalité des étudiants provenant des milieux défavorisés. Par exemple, une plus grande centralisation, au plan provincial, des programmes d'études entraîne inévitablement une plus grande uniformisation de ces programmes, laquelle est le plus souvent faite en fonction de normes implicites qui correspondent aux valeurs de certains milieux. Ces derniers ne sont habituellement pas les milieux défavorisés.

Les partis pris potentiellement les plus pénalisants devraient découler du contexte dont la pratique ne constitue qu'un reflet ou une conséquence. Nous mettrons pourtant l'accent sur les pratiques d'évaluation, ne considérant le contexte que pour bien dégager la position et l'importance relative des pratiques.

1. L'intelligence et l'apprentissage en tant qu'objets d'évaluation

Parmi les principaux objets d'évaluation que l'on retrouve dans le système scolaire (intelligence, apprentissage, aptitude, intérêts et personnalité, développement socio-affectif, développement psychomoteur, enseignement, programme, institution), l'intelligence et l'apprentissage constituent les risques de pénalisation les plus grands parce qu'ils donnent lieu aux pratiques les plus systématiques et les plus généralisées. En effet, tant les évaluations des aptitudes, des intérêts et de la personnalité que celles du développement socio-affectif ou du développement psycho-moteur sont réalisées sur une base trop individuelle et singulière pour connaître un impact déterminant sur les milieux défavorisés. De plus, lorsqu'elles sont utilisées par une commission scolaire, de

telles mesures n'apparaissent pas au dossier cumulatif d'un élève et ne sont qu'exceptionnellement considérées pour fin de classement ou de sélection.

Nous avons donc choisi d'analyser attentivement le contenu de certains instruments du ministère de l'Éducation du Québec (MEQ) et de la Commission des écoles catholiques de Montréal (CECM) qui portent précisément sur l'intelligence et l'apprentissage. Ces instruments, largement utilisés au Québec, contribuent fréquemment à définir la réussite ou l'échec scolaire et dans une proportion qui relègue au second plan la part de l'évaluation faite par l'enseignant.

1.1 L'intelligence en tant qu'objet d'évaluation

Tout en étant dénoncé par certains auteurs (Lawler, Tort) qui le considèrent bourgeois, le concept même d'intelligence continue d'occuper une large place et présente, à ce titre, un intérêt en tant qu'objet d'évaluation. Néanmoins, force est de constater que certaines définitions ou opérationnalisations du concept de l'intelligence (ou du fonctionnement intellectuel) sont le plus souvent trop influencées par les valeurs des groupes culturels dominants. Dans de tels cas, les tests d'intelligence (ou de fonctionnement intellectuel) risquent de prendre l'allure de tests d'adaptation ou d'intégration aux valeurs culturelles dominantes, et défavorisent par le fait même certains groupes plutôt qu'ils ne donnent, ainsi qu'on serait en droit de s'y attendre, un aperçu du niveau de fonctionnement cognitif d'un individu. Les nombreuses querelles entre chercheurs (Jensen, Bergsten-Brucefors, Burt, Dobzhansky, Lawler, Tort) concernant l'importance des influences respectives de l'hérédité et de l'environnement sur le fonctionnement intellectuel illustrent très bien ce phénomène. D'ailleurs, on a essayé de rendre les tests d'intelligence moins injustes pour les groupes minoritaires avec les tests «culture free» sans obtenir de résultats satisfaisants; c'est pourquoi on s'oriente maintenant vers le développement de tests d'intelligence «culture adjusted» qui tiennent compte des différences culturelles. Le «System of Multicultural Pluralistic Assessment» (SOMPA) de Jane Mercer est un exemple de ce type de test. Signalons aussi le BITCH (Black Intelligence Test of Cultural Homogeneity) qui a été développé en utilisant le langage des quartiers noirs, langage plus accessible aux enfants de ces quartiers; comme il fallait s'y attendre, les enfants noirs obtiennent de meilleurs résultats avec ce test d'intelligence. On constate donc qu'une mesure de l'intelligence ne saurait être parfaitement neutre et objective et qu'il faut, lors de l'élaboration d'un instrument de mesure, tenir compte des interférences culturelles. Mais même à cela, une «coloration culturelle» sera toujours présente, coloration dont il faut être conscient au moment d'interpréter les résultats de façon à en atténuer les effets.

Il ressort donc de cette première analyse que l'utilisation de mesures de l'intelligence peut contribuer à pénaliser systématiquement les milieux défavorisés: Qu'en est-il dans les faits? Une analyse rapide des divers instruments mis à la disposition des commissions scolaires par le MEQ, et il en va de même des tests de la CECM, illustre bien que l'écueil culturel n'a pas été évité. Nous prendrons, comme exemple, le test de fonctionnement intellectuel du MEQ à l'adresse des premières

années (6 ans) mais nous aurions pu tout aussi bien arrêter notre choix sur n'importe lequel des huit tests élaborés par le MEQ pour couvrir les différents niveaux du cours primaire.

Ce test comporte cinq sous-tests : 1 — vocabulaire, 2 — similitudes, 3 — classification, 4 — habileté numérique, 5 — mémoire. Douze items composent le sous-test vocabulaire, la description d'un objet étant donnée à l'enfant qu'il doit identifier parmi cinq illustrations. Ainsi, on dira : « Fais une grande croix sur l'image qui représente l'animal qui vit dans la mer et qui est le plus gros de tous les animaux ». L'enfant doit alors choisir parmi les dessins suivants : un éléphant, une girafe, une baleine, un phoque et un léopard. Tout de suite, on est à même de réaliser que l'enfant qui aura eu le privilège de trouver chez lui des livres de contes ou d'illustrations sur les animaux etc. (que la télévision ne saurait remplacer) sera nettement plus avantage que celui qui n'aura jamais vu un livre avant d'arriver à l'école. Autre exemple : « Fais une croix sur l'image qui représente un objet décoratif que l'on suspend à un mur ». L'enfant doit alors choisir entre le dessin d'une moquette, d'une peinture (ou cadre !), d'une cheminée (ou foyer), d'une lampe de bureau et d'une prise de courant (fiche murale). Ici, on le constate, la formulation de la question est quelque peu compliquée. Si l'objectif de l'item est de savoir dans quelle mesure un enfant est capable de reconnaître l'objet « cadre » ou l'objet « peinture », objectif déjà discutable parce qu'il s'agit d'un objet dont la présence est beaucoup plus fréquente en milieu favorisé que défavorisé, alors il aurait fallu trouver une combinaison de mots autre que « objet décoratif ». On peut supposer que le mot objet soit à peu près accessible à tous, encore que le mot « chose » aurait été plus près du langage populaire, mais dans le cas de l'adjectif « décoratif », on peut vraiment craindre qu'il ne soit beaucoup plus familier aux enfants issus de milieux bourgeois.

Au delà de cette interrogation, on peut même se demander pourquoi on fait appel à une périphrase plutôt que d'employer le terme propre, baleine dans le premier exemple et toile ou peinture ou cadre dans le second exemple. Deux hypothèses retiennent l'attention : ou bien on a voulu faire porter l'épreuve sur plus d'un mot, « objet décoratif » par exemple et le biais culturel n'est pas évité, ou bien on a voulu éviter la difficulté du choix d'un mot qui pourrait ne pas être ni le seul, ni le meilleur (vg toile, peinture ou cadre) et risquerait, de ce fait, d'introduire une difficulté non pertinente à l'objet de mesure. Mais alors, le résultat n'est guère mieux puisque l'on a trouvé une périphrase qui véhicule une difficulté analogue. On peut aussi ajouter que dans le sous-test de vocabulaire du test de préparation scolaire à la première année (test également considéré comme centré sur le fonctionnement intellectuel) on utilise des mots, puisque cette fois il ne s'agit pas de périphrases, qui versent de façon encore plus marquée dans le biais culturel : tente, bouilloire (pourquoi pas « canard » ou « bombe »), lampe, veston (pourquoi pas « coat » ou « jacket »), vadrouille (pourquoi pas « moppe »), bibliothèque (combien d'enfants de milieux défavorisés connaissent cette pièce de mobilier ?), nid d'oiseau, télescope (cet objet est-il très connu des enfants

pauvres ?), lampe de poche (pourquoi pas « flash-light ») et pneu (pourquoi pas « tire »).

Somme toute, on constate que dans presque tous les cas d'items de vocabulaire dont le but devrait être, à ce niveau, de s'assurer si tel ou tel objet est « nommé » par l'enfant, ou bien on a choisi un mot (habituellement « correct »...) qui favorise la classe culturellement dominante, ou bien on a sélectionné des objets dont le degré de familiarité pour l'enfant varie selon son origine socio-culturelle, ou bien on a fait les deux à la fois. Les sous-tests de vocabulaire ne semblent donc pas avoir échappé à la difficulté de la saturation culturelle.

Le sous-test des similitudes fait appel à la discrimination visuelle. En effet, parmi une collection de cinq dessins semblables mais différents par un détail, il doit reconnaître le dessin jumeau d'un étalon placé en marge. Sans doute, ce sous-test n'est-il pas aussi manifestement « coloré » sur le plan culturel que le précédent. Toutefois, il ne faudrait pas croire qu'il soit totalement neutre car l'enfant sera avantagé qui aura eu accès, dans sa famille, à diverses publications du type livre à colorier et à bricolage (Pif gadget, Journal Pilote et Spirou etc.) où l'on invite l'enfant à retrouver une forme en suivant des chiffres, à parcourir des labyrinthes, à découvrir les anomalies dans un dessin ou à identifier les différences entre deux dessins. Or, ce n'est pas dans les familles de milieux défavorisés que l'on retrouve le plus souvent ce matériel « éducatif ».

Le sous-test de classification exige de l'enfant qu'il reconnaisse, parmi cinq objets dessinés, celui qui ne va pas avec les autres. Par exemple, le chien ne va pas avec le stylo, le livre, le pupitre et la règle. Aux remarques formulées à l'endroit du sous-test précédent et qui s'appliquent également ici, il faut ajouter que le choix des objets, alors qu'il est souvent acceptable, verse parfois dans un certain ésotérisme, au moins pour les enfants de milieux défavorisés, et ce d'autant plus que les illustrations sont généralement d'assez mauvaise qualité. Exemple : l'enfant doit découvrir l'objet qui ne convient pas entre un violon (ou violoncelle), un tambour, un piano, une harpe (instrument de musique très peu répandu en milieu populaire) et un lutrin. La bonne réponse, c'est bien sûr le lutrin, qui n'est pas un instrument de musique, mais l'enfant qui, par hypothèse, n'aurait jamais vu une harpe saura-t-il choisir entre celle-ci et le lutrin ? Et ce n'est pas nécessairement son fonctionnement intellectuel qui est en cause.

Le sous-test d'habileté numérique comporte plusieurs items à contenu varié dont voici quelques cas types : 1 — Devant le dessin de 18 balles, on demande à l'enfant de faire « des croix sur 13 balles ». 2 — Devant cinq cadres regroupant respectivement les dessins de 7, 8, 9, 8 et 6 flèches, on demande à l'enfant de faire « des grandes croix dans les deux cadres qui contiennent le même nombre de flèches ». 3 — Devant cinq cadres regroupant respectivement les dessins de 7, 5, 8, 10 et 6 blocs, on dit à l'enfant : « Tu as 2 blocs et je t'en donne 5 autres, fais une grande croix dans le cadre qui contient le nombre de blocs que tu as maintenant ». 4 — De façon analogue et devant des illustrations de ballons, on dit à l'enfant : « Cinq ballons sont gonflés, tu en fais éclater deux. Fais une grande croix dans le cadre qui contient les ballons qui sont encore

glonflés ». 5 — Et encore : « Tu coupes deux pommes à moitié. Fais une croix dans le cadre qui contient les morceaux que tu as obtenus ». 6 — Autre type de problème : « Tu as dix sous et tu veux acheter des billes. Chaque bille compte 2 sous. Fais une croix dans le cadre qui contient le nombre de billes que tu peux acheter avec 10 sous ». 7 — Et enfin : « J'ai douze crayons. J'en donne 3 à Lise et 1 à Richard. Fais une grande croix dans le cadre qui contient le nombre de crayons qui me restent maintenant ».

On constatera que l'on demande à l'enfant de faire montre de ses capacités « arithmétiques », en partant de la numération, en passant par l'addition, la soustraction pour atteindre même le cas de la multiplication ou celui de problèmes impliquant plus d'une opération. Pour bien saisir la portée de ce sous-test il faut se rappeler qu'il s'adresse à des enfants qui sont en première année et dont les apprentissages scolaires en mathématiques ont à peine débuté. Or, on ne saurait prétendre que les questions posées aux enfants rejoignent un « potentiel mathématique » indépendant de tout apprentissage (social ? familial ? scolaire ?) concernant l'univers des quantités et des nombres. En effet, les opérations retenues dans l'épreuve sont trop complexes pour relever exclusivement du « nombre intuitif », concept numérique caractéristique des enfants de 6 ans et moins. Comme il risque d'arriver plus souvent, en milieu favorisé, que l'on apprenne les chiffres aux enfants à la maison ou qu'on leur confie de l'argent de poche etc., il s'ensuit qu'une fois de plus, on mesure davantage la qualité d'expérience et d'apprentissage socio-culturel d'un enfant que son niveau relatif de fonctionnement intellectuel.

Enfin, le sous-test mémoire, qui ne concerne que la mémoire immédiate, rejoint la capacité qu'a l'enfant de retenir de mémoire les 3, 4, 5, 6 ou 7 objets (selon le degré de difficulté de l'item) nommés par l'enseignant(e) et qu'il doit identifier d'une croix parmi des ensembles variant entre 7 et 12 dessins d'objets différents. Ainsi, on va demander à l'enfant de faire « une croix sur la montre, l'aquarium et le pinceau » ou « sur la hutte, le pneu et la hache » ou « sur le lapin, la bague, le tabouret et le pot à lait ». Deux remarques s'imposent ici. En premier lieu, il faut rappeler que la mémoire est une faculté qui se développe par l'exercice. Or, il est clairement établi que la mémoire verbale n'est pas également sollicitée dans les classes favorisées ou défavorisées parce que la communication verbale est beaucoup moins valorisée chez ces dernières. En second lieu, le choix des mots à mémoriser est tel qu'il joue vraisemblablement contre les enfants de milieux défavorisés. Selon toute probabilité, les termes aquarium, hutte, pneu, tabouret, pot à lait etc. ne sont pas également répandus dans toutes les classes de la société ; ainsi, l'enfant qui ne connaîtrait pas un mot, par opposition à celui qui le connaît, aura non seulement à le mémoriser, ce qui est déjà plus difficile parce qu'il s'agit d'un mot nouveau, mais encore à l'associer à l'un quelconque des objets dessinés faisant partie de l'ensemble. La tâche exigée n'est donc pas la même dans les deux cas et, de plus, on doit souligner que, parfois, le nombre de mots et/ou objets dessinés inconnus est trop élevé pour qu'une réponse ne soit donnée autrement qu'au hasard.

En somme, le contenu de ce test de fonctionnement intellectuel, mais on pourrait conclure de la même façon avec tous les autres, donne à penser qu'il est très largement contaminé ou saturé sur le plan culturel en faveur des classes aisées. Et il n'a pas été fait mention de la situation d'administration de l'épreuve non plus que de son caractère abstrait (test de type « papier-crayon ») qui jouent encore très nettement en faveur des milieux intellectuels. Pourtant, ce test, comme les autres de la même catégorie, constitue un excellent prédicteur du rendement scolaire. Faut-il s'en étonner ? À vrai dire, il n'y a rien là de plus normal. L'école vise précisément à faire réaliser aux enfants ces apprentissages de pensée abstraite. Il va donc de soi qu'un enfant dont l'intérêt a très tôt été éveillé pour la réalité intellectuelle marque non seulement une avance mais éventuellement une supériorité définitive sur celui qui n'a pas vécu cette expérience. Les tests de fonctionnement intellectuel portent sur ce « bagage » ; il est donc raisonnable de penser qu'ils mettent en lumière la supériorité des enfants issus d'un milieu propice et annoncent la probabilité qu'ils auront de conserver leur supériorité, voire même de l'accroître, via l'école qui va directement en favoriser le développement.

À ce titre, ces tests remplissent très bien leur fonction, mais on ne saurait prétendre qu'ils constituent une mesure du fonctionnement intellectuel en soi : ils mesurent un type donné de fonctionnement intellectuel dans un contexte très particulier. À noter que l'idée n'est pas de nier le lien entre le développement intellectuel et les occasions antérieures d'apprentissage. Toutefois, il nous apparaît important de signaler que certains types d'environnement culturel facilitent le type de développement intellectuel que l'on privilégie dans les tests du M.E.Q. ou, réciproquement, que les situations d'apprentissage choisies comme indicatrices de bon fonctionnement intellectuel avantagent des enfants qui ont eu un tel type particulier d'occasions d'apprentissage.

1.2 — *Les apprentissages en tant qu'objet d'évaluation*

Il est normal que l'on évalue les produits de l'éducation et que l'on établisse si les objectifs ont été atteints. Toutefois, il est possible que certains programmes traduits en objectifs d'apprentissage défavorisent systématiquement certains groupes. L'évaluation faisant partie intégrante du processus d'apprentissage, il est donc légitime d'utiliser l'apprentissage en tant qu'objet d'évaluation en se rappelant que la définition des programmes constitue un problème politique et non pas un problème d'évaluation et que certains désavantages, bien que révélés par l'évaluation, doivent plutôt être imputés à ces choix socio-politiques. Cependant, il demeure que les instruments d'évaluation peuvent accentuer ou non, atténuer ou pas ces désavantages ; aussi, allons-nous tenter de regarder plus avant de quoi il retourne.

Les évaluations d'apprentissage à caractère officiel diffèrent selon que l'on considère la situation qui prévaut à la CECM ou celle qui prévaut, semble-t-il, à travers la grande majorité des autres commissions scolaires de la Province. À la CECM, on dispose de tests de rendement en mathématique et en français pour chacun des niveaux

du primaire. Partout ailleurs, si l'on s'en remet à notre investigation, on ne dispose, pour tout test formel, que de ceux du MEQ. Or la liste de ces tests est la suivante :

CODE	TITRE	NIVEAU
130 010	Français, langue maternelle Compréhension de la langue écrite	3e primaire
130 030	Français, langue maternelle Compréhension de la langue orale	3e primaire
160 010	Français, langue maternelle Compréhension de la langue écrite	6e primaire
160 037	Français, langue maternelle Compréhension de la langue orale	6e primaire
160 217	Anglais, langue seconde Compréhension de la langue écrite	6e primaire
160 237	Anglais, langue seconde Compréhension de la langue orale	6e primaire
160 307	Mathématique	6e primaire

De façon générale, on peut donc affirmer, si l'on fait exception de la CECM, que la pénalisation via l'évaluation dont seraient victimes les milieux défavorisés est très circonscrite dans le temps et dans le nombre des épreuves. Nous examinerons donc les tests de français et de mathématique de 6^{ème} année primaire du MEQ et de la CECM.

Dans tous les cas, il s'agit de tests normatifs dont le contenu des items, sans les couvrir de façon exhaustive, est puisé à même les différentes notions des programmes-cadre du MEQ. Les questions sont nombreuses (entre 38 et 50) et touchent des domaines relativement variés dont la pertinence pour l'adaptation quotidienne de l'individu dans la société n'est pas toujours évidente. En français, par exemple, on interroge l'enfant sur le sens des préfixes et suffixes, sur le vocabulaire, sur les antonymes et les synonymes, le choix du terme propre etc. On ne perçoit pas, à travers ce test, la présence d'un bagage verbal minimal qui le sous-tendrait et qui pourrait constituer une sorte de seuil en deçà duquel un individu risquerait d'avoir de la difficulté à communiquer avec ses pairs. On se retrouve davantage devant un instrument qui, par la diversité de sa composition, s'avérera relativement discriminatoire et de ce fait bon prédicteur de la réussite scolaire au secondaire. Toutefois, la distance prise à l'endroit des acquis scolaires de base nous paraît trop grande pour que l'on puisse véritablement parler d'une évaluation des apprentissages : on utilise un vocabulaire recherché (vg « resquilleurs », « cordée » alpine), on cite un texte de Châteaubriand très éloigné de la réalité québécoise tant par la langue (« volets fermés ») que par le thème, et on décrit des comportements qui relèvent plus de la vie bourgeoise qu'ouvrière (« centres touristiques », passage aux douanes, scènes champêtres). Une composante culturelle est

donc présente qui souvent n'a même pas à voir avec l'orthographe, la syntaxe et l'analyse grammaticale qui, tout en étant également teintés culturellement, ont au moins le mérite de rejoindre la « mécanique » de la langue et l'enseignement qui est censé en être fait. À noter que le test du MEQ comporte beaucoup plus de questions de ce type « mécanique » (22 sur 45) que celui de la CECM (22 sur 38), mais la coloration culturelle y demeure importante. Finalement, aucune place n'est faite ni à l'expression, ni à la créativité, où chacune des différentes classes sociales pourrait peut-être davantage se retrouver encore qu'il serait difficile d'éviter l'émergence de « canons » de la qualité.

À propos des tests de mathématique, des remarques analogues s'imposent. Les questions y sont nombreuses et variées et, à l'instar des tests de français, parfois assez éloignées du quotidien. Ainsi, on trouve des questions sur l'associativité et la distributivité, sur le calcul dans les bases autres que 10, sur les axes de coordonnées, sur la théorie des ensembles etc. La saturation culturelle demeure donc puisque ces différents exemples illustrent bien l'orientation des épreuves vers des contenus à incidence sur les études supérieures plus que sur la maîtrise d'opérations mathématiques essentielles à l'insertion rapide sur le marché du travail. Par contre, les exemples utilisés, particulièrement dans le cas des problèmes raisonnés, font appel à un vécu relativement universel.

En résumé, il faut retenir de cette brève analyse des évaluations d'apprentissages qu'elles comportent un parti pris culturel, à un degré certainement moindre que les tests de fonctionnement intellectuel, mais quand même pénalisant pour les milieux défavorisés. En outre ces effets pénalisants, particuliers ou périphériques, risquent de s'ajouter à un désavantage de base sans doute plus important et déterminant qui vient des programmes dont les tests, rappelons-le, ne constituent qu'une forme de reflet ou de révélateur.

2. Les fins de l'évaluation

Les effets de désavantage à l'endroit d'une classe sociale donnée liés à l'évaluation sont relatifs aux fins auxquelles on la fait servir. En d'autres mots, les plus mauvais tests n'auront pas un effet négatif aussi direct si l'usage en est limité à l'analyse diagnostique ou à la recherche. Par contre, s'ils servent à la sélection ou au classement sélectif, leurs vices de fond prendront toute leur signification et joueront systématiquement contre les groupes d'individus par eux concernés. Au chapitre des fins de l'évaluation, on peut donc distinguer entre l'évaluation menée au service des apprentissages et celle menée au service des décisions administratives.

Une telle catégorisation suggère que certaines évaluations étudiantes ne sont pas au service de l'apprentissage même si le système scolaire existe en fonction de l'aide à l'apprentissage. Il devient important de ne pas confondre les évaluations qui sont directement au service de l'apprentissage avec celles qui sont indirectement au service de l'apprentissage. Par exemple, certains administrateurs scolaires pourraient nous

expliquer que les évaluations étudiantes pour fins de classification sont indirectement au service de l'apprentissage puisqu'elles visent, entre autres, à constituer des groupes tantôt homogènes, tantôt hétérogènes selon les orientations pédagogiques et susceptibles, ainsi réunis, de mieux apprendre. Selon notre classification, toutefois, il s'agit nettement d'une évaluation au service des décisions administratives même si ces décisions peuvent être indirectement au service des apprentissages. Il est évident que, sur le plan théorique, le pédagogique et l'administratif sont plutôt complémentaires ; toutefois, en pratique, le pédagogique risque d'être systématiquement subordonné à l'administratif dont relève la responsabilité de l'évaluation et qui, trop souvent, est coupé de la réalité pédagogique.

On retrouve dans la première catégorie, celle de l'évaluation au service des apprentissages, toutes les évaluations qui font partie intégrante de démarches directes d'aide à l'apprentissage telles l'évaluation pour fins :

- d'auto-analyse (incluant l'évaluation pour fins d'orientation)
- d'enseignement correctif
- de diagnostic

On retrouve dans la deuxième catégorie, celle de l'évaluation au service des décisions administratives, toutes les évaluations qui servent d'abord à la prise de décisions administratives telles que l'évaluation pour fins :

- de notation
- de certification
- de classification
- de sélection
- de contrôle régional ou provincial

On aura noté que l'évaluation pour fins de recherche n'apparaît pas dans cette classification ; de fait, elle peut aussi bien servir les fins pédagogiques que celles de l'administration suivant le contenu mais elle ne joue pas un rôle important dans la définition existentielle du désavantage scolaire des milieux défavorisés.

2.1 *L'évaluation au service des décisions administratives*

Il s'agit en définitive ici du type d'évaluation le plus délicat : ce sont les décisions administratives qui, à tous les échelons de l'institution éducative, viendront sanctionner par des faits les orientations économique-socio-culturelles avancées par la ou les classes sociales dominantes. Une pénalisation est souvent générée par l'utilisation que l'on fait d'une évaluation plus que par l'évaluation elle-même, comme en atteste l'analyse des pratiques scolaires décrites dans le document *L'organisation de l'enseignement préscolaire et primaire* (1978). Le ministère de l'Éducation, dans ses *Manuels d'information* et ses *Guides d'administration* à l'adresse des commissions scolaires et des enseignants et relatifs aux tests, fait des mises en garde très claires au moment de présenter les buts et les objectifs des tests de français et de mathématique pour éviter qu'on en fasse un usage abusif. Ainsi, on précisera que, « si le programme-cadre du MEQ a été le point de

départ pour déterminer les aspects à mesurer, le test porte néanmoins sur ce qu'on peut considérer comme prioritaire et qui devrait être atteint par la plupart des élèves, quel que soit l'enseignement offert. Le test garde un caractère indicatif de l'évolution de l'apprentissage dans le milieu ». Et plus loin, on stipule « qu'un test comme celui-ci n'est pas conçu pour mesurer l'atteinte de tous les objectifs d'un programme, *ni pour évaluer le rendement des enseignants* » (souligné dans le document). En effet, les objectifs que l'on fixe à l'usage des tests sont les suivants : « 1 — L'évaluation de la qualité des apprentissages, 2 — la poursuite d'une action pédagogique mieux éclairée et 3 — la sélection des anomalies, blocages et déficiences individuels ou collectifs des élèves ». Et l'on ajoute enfin, dans les « Guides d'administration », que « pour un élève il ne saurait être question d'échec ou de réussite au test et que, de plus, les résultats obtenus ne doivent pas par eux-mêmes servir à décider de sa promotion ». L'analyse du document précité indique que ces positions de principes ne sont pas respectées puisqu'effectivement les tests du MEQ sont vraisemblablement utilisés par 16 commissions scolaires sur 39 afin d'évaluer le « degré de préparation » des élèves pour le secondaire et procéder à leur classement dans des voies académiques différentes.

En somme, les concepteurs de tests peuvent faire le meilleur des travaux et donner tous les avertissements jugés nécessaires, rien pour autant ne saurait empêcher les usagers de poursuivre les fins de leur choix. Il ne faudrait donc pas confondre la portée et la signification des résultats d'évaluation avec la portée et la signification des décisions administratives tout comme il serait erroné de croire que la qualité scientifique d'un instrument d'évaluation garantit son caractère de neutralité ou d'objectivité. Mais il y a plus car, même dans l'hypothèse d'un test parfait, le recours à l'évaluation ne saurait assurer l'équité d'une sélection. À titre d'illustration, et peu importe si la situation évoquée doit sembler caricaturale, on peut se demander s'il est juste de procéder à quelque sélection, depuis le préscolaire jusqu'à l'universitaire, qui ne soit pas fondée sur les intérêts ou l'autodétermination. D'un strict point de vue théorique, l'idée n'est pas déraisonnable qui voudrait que l'on adapte ou individualise les enseignements à chacun des niveaux d'étude pour que tout individu parvienne au sommet de son potentiel, dut-il prendre deux fois le temps d'un autre plus favorisé. En pareil cas, toute sélection basée sur une évaluation normative est inéquitable par définition, peu importe sa valeur intrinsèque.

Ajoutons, enfin, que le recours à l'évaluation normative est tout à fait inutile au niveau primaire puisque le MEQ y a décrété la règle de la promotion automatique. Il ne devrait donc y avoir ni classement ni sélection nécessaires ; aussi, devrait-on encourager les commissions scolaires à se centrer sur la démarche de l'apprentissage beaucoup plus que sur son aboutissement et, à fortiori, que sur son aboutissement comparé.

2.2 *L'évaluation au service des apprentissages*

Le texte qui précède laisse déjà clairement entrevoir que les désavantages liés à l'évaluation pourraient être grandement réduits si l'on se centrait sur le service à

l'apprentissage plutôt que sur la comparaison des résultats une fois la démarche terminée. En fait, il est difficile d'imaginer qu'une évaluation dont le but premier serait d'aider à mieux intervenir dans un processus d'apprentissage puisse présenter des risques de pénalisation systématique aussi importants. Et même dans l'hypothèse où la chose pourrait se produire, il s'agirait alors d'une pénalisation par défaut et non d'une pénalisation active dans le sens où la conséquence ne serait pas tant de déclasser un individu que de ne pas favoriser une intervention aussi efficace dans sa démarche d'apprentissage.

3. Les approches d'évaluation

Nous nous intéresserons à deux approches d'évaluation : l'approche normative et l'approche critériée.

Dans le cas de *l'évaluation normative*, les instruments sont élaborés de façon à amplifier les différences entre les divers répondants. Conséquemment, l'interprétation des résultats d'un répondant ne peut se faire qu'en fonction des résultats des autres répondants. Aussi, lors de l'élaboration d'un test de rendement *normatif*, le constructeur de tests choisira les items (questions) qui contribueront le plus à maximiser les différences entre les divers répondants tout en mesurant l'objet d'évaluation visé. Une telle pratique a pour effet de sacrifier partiellement la validité de contenu d'un test au profit de sa puissance de discrimination entre les répondants. En effet, les parties du domaine mesuré qui génèrent peu de différences entre les répondants risquent d'être sous-représentées alors que les parties qui différencient beaucoup entre les répondants risquent d'être surreprésentées.

De plus, lors de l'interprétation des résultats obtenus à un test de rendement normatif, la capacité d'inférer au domaine mesuré est sacrifiée à la capacité de différencier le mieux possible entre les répondants et de les situer les uns par rapport aux autres. Il est très important de savoir que les résultats obtenus à un test de rendement normatif permettent *d'abord* et *surtout* de situer le répondant par rapport aux autres répondants. Évidemment, le résultat obtenu à un test normatif fournit de l'information quant à la maîtrise du domaine mesuré mais la valeur de cette information dépend de la validité de contenu du test qui, dans un tel cas, est subordonnée à la puissance discriminatoire du test.

Dans le cas de *l'évaluation critériée*, les instruments sont élaborés de façon à pouvoir situer le mieux possible chaque répondant par rapport au domaine mesuré, c'est-à-dire par rapport à l'ensemble d'objectifs à atteindre lors d'une séquence d'apprentissage s'il s'agit d'un test de rendement. Conséquemment, l'interprétation des résultats d'un répondant ne peut se faire qu'en fonction du domaine mesuré. Ainsi, lors de l'élaboration d'un test de rendement critérié, le constructeur de test choisira les items (questions) qui représentent le mieux possible les différentes parties du domaine mesuré sans se préoccuper de la capacité de différencier entre les répondants. Une telle pratique a pour effet de sacrifier partiellement la puissance discriminatoire d'un test au

profit de la validité de contenu. En effet, toutes les parties du domaine mesuré sont adéquatement représentées dans le test, ce qui permet de considérer l'ensemble des questions du test critérié comme un échantillon représentatif de l'univers des éléments du domaine mesuré. Il est très important de savoir que le résultat obtenu à un test de rendement critérié permet *d'abord* et *surtout* de situer le répondant par rapport au domaine mesuré, c'est-à-dire l'ensemble d'objectifs visés. Évidemment, le résultat obtenu par un répondant peut être comparé au résultat obtenu par un autre répondant pour savoir si le premier répondant maîtrise mieux tel type d'objectifs que le second répondant. Toutefois, il faut se rappeler que le test critérié ne vise en aucune façon à amplifier les différences entre les répondants. Il ne faudra donc pas se surprendre de constater que certaines questions ont été réussies par tous les répondants alors que d'autres questions ont été échouées par tous les répondants. De telles questions qui ne contribuent aucunement à la puissance discriminatoire d'un test auraient été éliminées lors de l'élaboration d'un test normatif.

De plus, il faut comprendre que les approches normative et critériée peuvent être toutes deux utilisées pour élaborer des tests qui serviront tant à l'évaluation formative qu'à l'évaluation sommative. En effet, même si les tests critériés sont généralement associés à l'évaluation formative, le cumul de l'information ainsi recueillie rend possible une forme d'évaluation sommative. Il en va de même avec les tests normatifs qui, bien que plus généralement associés à l'évaluation sommative, peuvent être utiles dans l'évaluation formative. Il faut cependant réaliser que l'initiative en est alors laissée à l'enseignant qui doit, au delà des résultats comparatifs de moyenne ou de pourcentage de réussite, se pencher sur le contenu du test de façon à situer chacun de ses élèves dans son cheminement disciplinaire personnel. Il s'agit donc ici d'une utilisation qui n'est pas facilitée par le test et qui suppose un effort particulier de la part de l'usager.

Suite aux descriptions faites précédemment des approches normative et critériée, il se dégage clairement que l'approche normative est la plus susceptible de pénaliser systématiquement les étudiants des milieux défavorisés. En effet, cette approche vise à accentuer les différences entre les étudiants et, ce faisant, elle risque d'accroître la marginalité des étudiants des milieux défavorisés. Il ne faut toutefois pas s'imaginer que l'approche critériée règle tous les problèmes, mais il est permis de croire qu'elle est moins désavantageuse. Il faut se rappeler que les approches d'évaluation sont utilisées pour mesurer des objets précis d'évaluation dans des contextes précis d'évaluation et que, bien utilisées, elles ne peuvent que refléter les options inhérentes aux objets et aux contextes d'évaluation. Ainsi, l'utilisation de tests critériés pour évaluer l'apprentissage dans une matière définie selon les normes d'un milieu précis ne corrigera pas les aspects culturels inhérents au programme en question. Toutefois, l'utilisation de cette approche, puisque centrée sur le cheminement d'un individu dans son apprentissage, n'ajoutera pas à l'effet du parti pris culturel. Par contre, l'utilisation de l'approche normative pourra contribuer à son influence sur les résultats des étudiants car la maximisation des différences signifie aussi maximisation, entre autres, des différences à

teneur culturelle. De plus, l'utilisation de l'approche normative pour évaluer l'apprentissage peut facilement engendrer un processus de sélection continue qui risque de pénaliser les milieux défavorisés ; l'utilisation de l'approche normative garantit le maintien des différences individuelles au niveau de l'apprentissage puisqu'elle garantit qu'il y aura toujours des meilleurs et des pires, c'est-à-dire qu'elle assure l'inégalité des résultats. Bien qu'il faille distinguer entre égalité des résultats et égalité des chances, nous estimons que cette pratique d'évaluation cadre mal dans l'application d'une politique d'égalisation des chances en éducation. Il nous semble donc que :

- 1) L'utilisation de l'approche normative est la plus susceptible de pénaliser systématiquement les étudiants des milieux défavorisés.
- 2) L'utilisation de l'approche critériée devrait moins contribuer à amplifier la marginalité des milieux défavorisés.

Par ailleurs, un rapide coup d'œil sur la situation actuelle de l'évaluation dans les commissions scolaires permet de constater que l'approche normative est très répandue alors que l'approche critériée n'est que très peu utilisée, même si cette dernière est beaucoup mieux adaptée à une approche pédagogique par objectifs. En rapport avec cette pédagogie par objectifs, il faut toutefois souligner que, pour la mesure de l'apprentissage, on tend à adapter l'approche normative de façon que les objectifs pédagogiques y soient mieux représentés. Dans cette approche normative adaptée, on sélectionne toujours les questions qui amplifient les différences entre les répondants ; cependant, cette sélection de questions se fait à l'intérieur de chaque partie du domaine à évaluer et assure ainsi, sur le plan de la structure, une représentativité de base. Mais toujours, on se retrouve avec les questions qui vont accentuer les différences entre les répondants. Cette adaptation présente donc le mérite d'assurer une validité de contenu minimale mais ne corrige en aucune façon les risques de pénalisation systématique pour les étudiants provenant de milieux défavorisés.

4. Pénalisation qui résulte d'une centralisation pédagogique au niveau des programmes d'études

Une centralisation des programmes d'études signifie que les mêmes programmes seront étudiés dans les diverses régions de la province ; il s'agit donc d'une démarche qui uniformise le système d'éducation et qui réduit les différences régionales. Une telle centralisation permet, entre autres, une plus grande mobilité entre les diverses régions scolaires, c'est-à-dire qu'un enfant qui passe d'une région à une autre continue à étudier les mêmes programmes. Une telle centralisation permet aussi de réaliser plus facilement des études comparatives entre les régions et entraîne inévitablement un plus haut niveau de contrôle gouvernemental sur les produits de l'éducation. Il va de soi que le niveau de centralisation est plus élevé lorsque les objectifs d'un programme provincial sont très précis et qu'une telle situation laisse peu de place aux variations et adaptations régionales ou locales. En conséquence, le rôle d'évaluation des instances régionales ou locales se trouve réduit lorsque les objectifs provinciaux d'un

programme sont très précis. Par ailleurs, étant donné que les décisions concernant l'importance relative des matières et les contenus des matières constituent des décisions « politiques », il s'ensuit que les milieux qui influencent le plus ces décisions politiques risquent d'être systématiquement favorisés tandis que les milieux qui influencent le moins ces décisions politiques risquent d'être systématiquement défavorisés. Par exemple, si un programme hautement centralisé est défini en fonction des valeurs d'un milieu « x », la probabilité est grande que les enfants du milieu « x » aient systématiquement plus de facilité à assimiler ce programme que les enfants provenant d'un milieu « y » dont les valeurs sont différentes, et qui, de ce fait, risquent d'être systématiquement désavantagés dans l'apprentissage des objectifs de ce programme. On peut donc croire que, dans une société hétérogène, toute centralisation des programmes d'éducation favorisera certains groupes et en défavorisera d'autres. Suite à ces précisions, nous voyons que ce type de pénalisation est directement et principalement lié aux processus politiques de définition des objectifs des programmes et que les types d'évaluation étudiés précédemment ont ici un rôle périphérique. En effet, tel que mentionné antérieurement, les procédures et instruments d'évaluation devront respecter les contenus des programmes et ne pourront, s'il s'agit de bonnes procédures et de bons instruments, que refléter les partis pris inhérents aux programmes. Rappelons encore ce que nous avons déjà dit au sujet de l'approche normative, à savoir qu'elle risque de contribuer à amplifier l'effet des options culturelles inhérentes aux programmes sur les résultats des étudiants. Il ne faudrait donc pas s'attendre à ce que les procédures et instruments d'évaluation corrigent ces aspects ; une démarche correctrice basée sur de telles attentes risque d'aboutir à des prescriptions qui ne s'intéressent qu'aux symptômes plutôt qu'à la cause des différences observées.

Sans prétendre à une analyse systématique des programmes d'études primaires dans notre système d'éducation, il est facile de constater qu'ils manifestent une orientation implicite vers la préparation aux études supérieures (universitaires) plutôt que vers la préparation au milieu du travail. L'introduction de la théorie des ensembles dès le niveau primaire nous semble une bonne illustration de cette orientation implicite. Semblable orientation implicite ne risque-t-elle pas de générer des programmes d'études primaires dont les contenus sont susceptibles de pénaliser systématiquement les étudiants des milieux défavorisés ? L'objectif de préparation au marché du travail ne devrait-il pas occuper une place plus grande dans les programmes d'études primaires ? Tout en reconnaissant l'importance de la préparation aux études supérieures, il faut bien admettre que cette place prioritaire est occupée au détriment d'autres apprentissages qui seraient potentiellement plus pertinents pour d'autres catégories d'étudiants, lesquels apprentissages seraient moins susceptibles d'introduire un désavantage systématique pour les étudiants provenant de milieux défavorisés, parce qu'ils sont plus près de leurs valeurs. Nos programmes d'études primaires préparent encore les étudiants à entrer au « collège classique » et le livre blanc sur l'École québécoise n'annonce pas de changement majeur à ce niveau. L'insistance sur l'apprentissage du français standard international constitue un autre type de contexte d'évaluation

éventuellement pénalisant pour les milieux défavorisés. Si, donc, l'on additionne les effets de l'approche normative aux différences initiales, on comprendra facilement que les étudiants provenant des milieux défavorisés risquent d'être ainsi doublement pénalisés.

Conclusion

Les implications pratiques de l'évaluation pour les milieux défavorisés, on l'aura noté, ne sauraient être négligées tant elles sont nombreuses et complexes, sans parler de leur variété souvent insoupçonnée. En résumé, rappelons que les tests d'intelligence, reflets d'une « performance sociale valorisée » risquent de pénaliser systématiquement les enfants de milieux défavorisés tout comme le font, même si c'est à un degré moindre au plan du contenu, les tests d'apprentissage, reflets des orientations données aux programmes pédagogiques. Dans les deux cas, on constate l'importance déterminante des « définitions » que se donnent les responsables de l'évaluation, les grands gestionnaires de l'institution scolaire et, finalement, la société toute entière, de ces concepts combien relatifs que sont l'intelligence et la réussite scolaire. Dans les deux cas également, on a observé que l'évaluation était de type normatif, qu'elle était, à cause de cela et aussi faute de poursuivre résolument d'autres fins, au service des décisions administratives beaucoup plus qu'au service direct des apprentissages. Il peut en résulter alors une maximisation des différences individuelles qui entraîne par définition la catégorisation d'enfants « faibles » en nombre pré-déterminé et favorise l'action de ce que, depuis Rosenthal, on appelle l'effet pygmalion, c'est-à-dire cette tendance que l'on a généralement à « voir » faible ou déficient un enfant déjà reconnu tel à tort ou à raison.

L'évaluation normative, la plus développée par le MEQ et toujours la plus répandue dans les commissions scolaires, malgré quelques efforts embryonnaires faits dans d'autres directions, contribue largement au maintien d'un statut qui n'aide en rien les milieux défavorisés. Par sa nature, elle a peu d'impact positif sur la démarche d'apprentissage, si ce n'est de manière très indirecte, et elle constitue une donnée de base dans les dossiers cumulatifs qui « suivent » et souvent « pénalisent » les enfants toute la durée de leurs études et même au delà. Elle inquiète bon nombre d'enseignants qui, dès lors, orientent leur enseignement en vue d'une réussite de leurs élèves aux tests et, enfin, elle sert souvent, malgré les directives contraires du MEQ, au classement au secondaire. Elle conserve une utilité pour la recherche et la gestion en fournissant des informations comparatives sur les différentes régions de la Province, mais gagnerait à être limitée à cette fin, ce qui supposerait une diffusion des tests et des résultats autrement régie qu'elle ne l'est présentement.

Quant aux programmes, on a réalisé que dans la série des causes potentielles de pénalités pour les milieux défavorisés, ils constituaient l'une des plus importantes, si ce n'est la plus importante. Même si ce devrait être la responsabilité des conseillers en mesure et évaluation de signaler la portée et les limites des instruments de mesure, il faut dire, à leur décharge, que l'évaluation ne fait que dégager les résultats sur les

enfants d'interventions « orientées ». En effet, l'évaluation ne saurait, au risque d'être un mauvais outil mais également parce qu'elle ne pourrait y parvenir absolument, compenser pour les biais inhérents à l'institution scolaire. Seule une prise de conscience collective des anomalies d'un système pourrait entraîner, à moyen et à long terme, la mise en place de mesures correctives appropriées, lesquelles devraient d'ailleurs déborder le cadre de l'école pour rejoindre le marché du travail et la question épineuse de la répartition des richesses. Entretemps, certes, on pourrait améliorer la situation en atténuant le caractère de préparation aux études supérieures des programmes du primaire et en procédant à des évaluations moins normatives par l'insistance un peu moins grande sur le français standard international, à titre d'exemple...

Les pratiques pédagogiques, à l'instar des programmes dont elles dépendent largement, accusent le même parti pris culturel, participent essentiellement des mêmes valeurs. Elles ne sauraient donc avantager les milieux défavorisés puisqu'elles versent le plus souvent dans une forme d'élitisme encouragé et accentué par l'évaluation normative.

Somme toute, la question n'est pas seulement, ni même principalement, une question d'évaluation. La pénalisation qui pèse sur les milieux défavorisés est bien plus complexe et ne saurait être ainsi réduite.

Néanmoins, avec la promotion automatique, le terrain est tout indiqué pour œuvrer à l'application d'une conception de l'évaluation qui pourrait précéder ou accompagner l'implantation de nouveaux programmes et de nouvelles pratiques pédagogiques. Avec la publication du livre blanc sur l'éducation, il ne s'agit peut-être pas tant de rechercher les bouleversements complets mais plutôt l'évolution progressive des mentalités, laquelle serait sans doute aidée, en milieu scolaire, par l'élaboration et la mise à disposition de nouveaux types d'instruments et de nouvelles orientations en évaluation.

BIBLIOGRAPHIE

- Angers, P. (1976), *Les modèles de l'institution scolaire*. Centre de développement en environnement scolaire. Contribution à l'analyse institutionnelle.
- Bergsten-Brucefors, Agneta (1976), *Intelligence patterns and their Relation to Social Background*, Lund : CWK GLEERUP.
- Bloom, B.S., Hastings, G.T. et Madaus, G.F. (1971), *Handbook on Formative and Summative Evaluation of Student Learning*.
- Burt, C. (1966), *The Genetic Determination of Differences in Intelligence*, *The British Journal of Psychology*, 57, 137-153.
- Dobzhansky, T. (1973), *Genetic Diversity and Human Equality* New-York : Basic Books, Inc.
- Jensen, A.R. (1972), *Genetics and Education*, New-York : Harper and Row.
- Lawler, J. (1978), *Intelligence, Génétique et Racisme*, Paris : Éditions Sociales.

- Lépine, G. (1977), *Analyse des modèles utilisés en éducation au Québec*. Montréal, Les Éditions coopératives Albert St-Martin.
- Popham, W.J. (1971), *Criterion — Referenced Measurement*, Prentice-Hall Inc., Englewood Cliffs, New-Jersey.
- Rosenthal, R., and Jacobson, L. (1968), *Pygmalion in the Classroom*, New-York : Holt, Rinehart and Winston.
- Tort, M. (1914), *Le Quotien intellectuel*, Paris, François Maspero.
- Trahan, M. et Dassa, C., *La Courbe normale en psychométrie et en éduométrie : postulat, constante naturelle ou erreur scientifique*. Revue des sciences de l'éducation, vol. IV, no 1, hiver 1978.