

# Mesurer la compétence à écrire : mise à l'essai de prototypes d'items à réponse choisie portant sur la cohésion sémantique

Jean-Guy Blais and Christian Rousseau

Volume 18, Number 1, 1995

URI: <https://id.erudit.org/iderudit/1092489ar>

DOI: <https://doi.org/10.7202/1092489ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Blais, J.-G. & Rousseau, C. (1995). Mesurer la compétence à écrire : mise à l'essai de prototypes d'items à réponse choisie portant sur la cohésion sémantique. *Mesure et évaluation en éducation*, 18(1), 59–94.  
<https://doi.org/10.7202/1092489ar>

Article abstract

The present study deals with the measurement of a specific aspect of writing competency. It is an exploratory study and its goal is to analyse the validity of interpretations done about writing competency with scores coming from answers to multiple-choice items. To do so, items prototypes were developed and tried with a small group of students at the college level. Each examinee score is compared to content analysis of one essay and is explained with a theoretical framework of writing competency. Even if our results are limited to the present study, they constitute an opening to study the use of multiple choice items to assess other aspects of writing competency than the linguistic code.

## Mesurer la compétence à écrire: mise à l'essai de prototypes d'items à réponse choisie portant sur la cohésion sémantique

Jean-Guy Blais  
Christian Rousseau  
Université de Montréal

*L'étude présentée s'intéresse à la mesure d'un aspect de la compétence à écrire, les macrostratégies de cohésion sémantique, à l'aide d'items à réponse choisie (questions à choix multiples). Elle est menée dans une perspective exploratoire. Le but de l'étude est d'analyser la validité des interprétations que les scores obtenus à des items à réponse choisie permettent de formuler à propos de la compétence à écrire. Pour ce faire, des prototypes d'items ont été mis à l'essai auprès d'un petit groupe d'élèves de première année du collégial. Le score de chacun des candidats est comparé à une analyse de contenu d'un texte produit par lui et est interprété en fonction d'une représentation théorique de la compétence à écrire. L'analyse s'intéresse à une qualité importante des interprétations formulées, soit leur validité nomologique. Les résultats, bien que limités à l'étude présentée, impliquent principalement une ouverture sur la possibilité de concevoir des items à réponse choisie pour mesurer d'autres aspects de la compétence à écrire que la maîtrise du code linguistique.  
(compétence à écrire, cohésion sémantique, items à réponse choisie)*

*The present study deals with the measurement of a specific aspect of writing competency. It is an exploratory study and its goal is to analyse the validity of interpretations done about writing competency with scores coming from answers to multiple-choice items. To do so, items prototypes were developed and tried with a small group of students at the college level. Each examinee score is compared to content analysis of one essay and is explained with a theoretical framework of writing competency. Even if our results are limited to the present study, they constitute an opening to study the use of multiple choice items to assess other aspects of writing competency than the linguistic code.*

*(writing competency, semantic cohesion, multiple choice items)*

## Introduction

La mise au point d'items pouvant fournir des données valides non plus uniquement sur la connaissance des aspects normatifs de la langue écrite, mais aussi sur des habiletés dites «supérieures» faisant partie de la compétence à écrire, demande l'élaboration de cadres de référence «globalisants». Ainsi, par exemple, une proposition de Frederiksen et Dominic (1981) suggère d'envisager l'écriture comme une activité cognitive, comme une forme particulière de langage et d'utilisation du langage, comme un processus communicatif et, enfin, comme une activité contextualisée, intentionnelle.

Évidemment, il est peu probable que, dans l'immédiat, on puisse facilement arriver à élaborer des instruments de mesure de la compétence à écrire dans la perspective englobante de Frederiksen et Dominic. Toutefois, en explorant les propriétés de certains types d'items, en accumulant des données sur leurs possibilités et leurs limites quant à la mesure de cette compétence complexe, il est permis de croire que certaines contributions peuvent être apportées en une matière où les besoins actuels de l'éducation sont nombreux. C'est dans cette direction que s'oriente cette recherche.

## Problématique

### Les besoins en matière d'instrumentation

La faiblesse des résultats à l'épreuve de rédaction administrée par le ministère de l'Éducation du Québec aux finissants du secondaire, en mai 1986 (Gouvernement du Québec, 1986), a sans doute constitué le point culminant de plusieurs signaux d'alarme au sujet de la qualité des textes écrits par les élèves. Ces résultats furent suivis par ceux d'une vaste enquête sur les perceptions et les attentes de la société québécoise face à l'enseignement du français, langue maternelle (Bibeau, Lessard, Paret & Thérien, 1987). Selon cette enquête:

D'après l'idée que se fait la bonne majorité des Québécois, les jeunes qui sortent des écoles secondaires, en majorité, ne savent ni bien parler, ni bien écrire, ni lire et n'ont pas développé le goût de la lecture: toutefois, ils savent écouter et comprendre ce que les autres disent.

Pour que la volonté de transformer la situation insatisfaisante de l'écriture des élèves québécois puisse se traduire en propositions

pertinentes, il importe de bien décrire cette situation. Or, que sait-on de la compétence des élèves en matière d'écriture? L'épreuve de rédaction permet de constater certaines caractéristiques des textes produits par les élèves, mais l'évaluation d'un texte unique, produit en réponse à une tâche d'écriture qui fixe le sujet, le type de texte et le contexte de production, permet-elle de décrire suffisamment la compétence à écrire pour les besoins de l'enseignement, du classement ou de la certification? Quels autres outils sont disponibles pour parvenir à ces fins?

Les cadres de référence officiels en matière d'écriture sont ceux que véhiculent les programmes du primaire et du secondaire et les descriptions de cours du collégial. Les formulations qu'on y retrouve, du type «Produire des textes ayant telles ou telles caractéristiques», associent la compétence à la performance. Les contenus abordés dans les programmes se rattachent donc aux caractéristiques souhaitées du produit textuel (allant des aspects normatifs de la langue aux aspects organisationnels plus complexes, en passant par les divers types de texte et modes de discours). Les dispositions relatives à la production comme telle (comment produire un texte?) sont à peu près absentes. Quant aux données disponibles, elles sont tributaires de ces cadres de référence et elles concernent d'abord les caractéristiques des rédactions d'élèves, comme c'est le cas de celles qui sont recueillies lors de l'épreuve ministérielle de fin de secondaire. Les autres données se rapportent à des aspects relativement restreints de la compétence à écrire (la connaissance de l'orthographe, de la ponctuation, du lexique ou de la syntaxe). C'est le cas des tests dits «objectifs» qu'administraient les universités francophones du Québec à leurs nouveaux candidats, par exemple, ou du *Test de classement en français écrit (TEFEC)* utilisé dans certains cégeps (Ferland & Sanchez, 1990).

Le manque de cadre de référence et de données n'empêche pas que des décisions politiques d'envergure soient prises. Au test d'entrée des universités, dès le printemps de 1992, se rajoutait une épreuve de rédaction, calquée sur l'épreuve de fin d'études secondaires, administrée par le ministère de l'Enseignement supérieur et de la Science de l'époque aux finissants du collégial se destinant aux études universitaires. Cette épreuve a soulevé de vives réactions dans le milieu collégial (Lecavalier, 1991; Asselin & Mc Laughlin, 1992).

Il nous apparaît, en bout de ligne, que les instruments de mesure dont on dispose actuellement semblent insuffisants pour recueillir efficacement des données pertinentes afin d'évaluer les différentes facettes de la compétence des élèves à produire des textes.



## Les procédures directes et indirectes de mesure

En général, autant pour l'écrit que pour d'autres matières, les situations qui exigent une production de la part de l'élève donnent lieu à deux procédures de collecte d'information. La première, caractérisée par le fait qu'elle utilise une production du candidat, est souvent nommée «directe», alors que la seconde, dite «indirecte», demande au candidat qu'il réponde à des questions portant sur son habileté à réaliser cette production.

Pour des raisons pratiques de communication, la présente étude conserve les qualificatifs «directes» et «indirectes» relativement aux procédures de mesure, même si leur emploi peut poser problème. En effet, même si ces termes se retrouvent dans de nombreuses études se rapportant à la mesure de la compétence à écrire, la logique qui permet d'interpréter un score obtenu à une tâche de mesure comme une mesure de la compétence procède toujours par inférence, donc indirectement. Il est impossible de mesurer «directement» une compétence. En réalité, lorsqu'on observe le produit (un texte) résultant de la compétence à écrire, c'est l'observation qui est directe et non la mesure.

Néanmoins, par procédure directe de mesure de la compétence à écrire, on entend l'attribution d'un score par un correcteur à un texte produit par un candidat, texte qu'on considère comme un représentant de tous les textes que pourrait produire cet individu. On peut obtenir ce score en portant un jugement sur l'ensemble du texte et en lui faisant correspondre une valeur sur une échelle prédéterminée (évaluation dite «holistique»). On peut aussi analyser le texte sous différents aspects (correspondant au fond et à la forme, par exemple) et accorder un score pour chacun de ces derniers (évaluation dite «analytique»). On peut, de là, calculer un total ou un total pondéré. C'est cette perspective qui est privilégiée lorsqu'on utilise une «grille d'appréciation» pour la correction.

Pour ce qui est des procédures indirectes de mesure de la compétence à écrire, elles ne requièrent pas d'échantillon de texte d'un candidat et peuvent être obtenues, par exemple, à l'aide d'items à réponse choisie.

Cette distinction entre «directe» et «indirecte», est préférée à «subjective» et «objective» pour désigner ces pratiques, d'abord parce qu'on peut obtenir une mesure objective à partir d'un texte (en comptant, par exemple, le nombre de mots ou le nombre de fautes d'orthographe) et ensuite parce qu'il est abusif de prétendre à l'objectivité d'un item à réponse choisie (sa correction est objective, mais la forme et le contenu de l'item demeurent des choix subjectifs).

La distinction entre procédures directes et indirectes est étroitement reliée au concept de compétence. Pour cette raison, Bachman (1990) parle de deux différentes perspectives dans la mesure de la compétence langagière: une approche «réalité» (*real-life*, procédure directe), à l'intérieur de laquelle la compétence est définie de façon «opérationniste», c'est-à-dire en fonction de la performance uniquement; et une approche «habileté/interaction» (*interactional/ability*, procédure indirecte), où la compétence est définie en fonction des habiletés sous-jacentes de production. Le choix d'un type de procédure de mesure ne devrait donc pas être arbitraire, mais dépendre d'une représentation conceptuelle de l'objet de la mesure.

### Les limites des procédures directes et des procédures indirectes

L'argumentation principale en faveur des procédures dites directes de mesure repose sur l'idée que la meilleure façon de connaître l'habileté d'écriture d'un candidat est de lui demander d'écrire un texte. Ces procédures jouissent, de ce fait, d'une certaine crédibilité (*face validity*) auprès des enseignants et du public en général. Dans le cadre d'une épreuve à grande échelle, le message implicite d'une épreuve de rédaction peut aussi entrer en ligne de compte. Ainsi, parce qu'elle valorise l'activité d'écriture, la rédaction motiverait les candidats à maîtriser les habiletés requises (Conlan, 1986). Cette façon de recueillir l'information suscite toutefois certaines interrogations sur le plan de la validité et de la fidélité.

La critique de la validité des procédures directes repose principalement sur le fait qu'on y confond l'habileté elle-même avec l'observation d'une production censée exiger cette habileté. Cette identification de l'habileté et de la performance ne permet pas de faire des inférences au-delà du contexte de mesure et limite, d'une part, l'interprétation et l'utilisation des résultats et, réciproquement, le type d'évidences qui peut les supporter (Bachman, 1990).

De plus, cette identification de l'habileté et de la performance entraîne à la fois l'avantage et l'inconvénient de permettre d'évacuer tout recours à une définition de la compétence à écrire. Comme on peut la définir à partir de son résultat (un texte), on croit souvent qu'il n'est pas nécessaire de s'encombrer d'une représentation théorique de cette compétence. Ce choix limite cependant la validation d'une procédure spécifique. Outre la crédibilité, qui repose surtout sur une apparence de validité — ce qui ne peut pas être considéré comme une évidence de validité —, la seule avenue praticable afin de procéder à une validation d'une procédure directe demeure de tenter de démontrer la représentativité du contenu de

l'épreuve. On doit alors montrer comment la tâche proposée constitue un échantillon d'une situation authentique d'écriture, ce qui est extrêmement compliqué, l'authenticité s'avérant la plupart du temps contestable (Bachman, 1990). On peut citer d'autres problèmes de validité soulevés par les procédures directes. Un candidat peut arriver, par exemple, à masquer ses faiblesses en utilisant des stratégies de détour ou d'évitement lorsqu'il ne maîtrise pas certaines habiletés qu'on souhaiterait pourtant voir chez lui (Stiggins, 1982). La mesure porte alors sur un texte qui ne représente que partiellement le contenu visé.

La lacune des procédures directes la plus souvent soulignée est leur manque de fidélité, ce qui constitue une limite supplémentaire de leur validité. En effet, même dans le cas d'une mesure dont les qualités de validité sont élevées selon les normes établies, le fait de fournir des données qui varient d'une prise de mesure à l'autre dans un moment déterminé pour le même objet (une certaine habileté chez un certain candidat, par exemple) restreint les interprétations qu'il est possible de faire à partir de ces mesures. Les résultats de nombreuses recherches (voir Breland, Camp, Jones, Morris & Rock, 1987) ont illustré que le score obtenu par les procédures directes de mesure de la compétence à écrire pouvait varier fortement d'une prise de mesure à l'autre et qu'il dépend, entre autres, du correcteur, du sujet du texte, du type de discours, du contexte de correction (après un bon ou un mauvais texte), de la qualité de l'écriture manuscrite ou de la propreté de la copie. Il existe des stratégies qui permettent de calibrer les correcteurs pour ensuite pondérer les scores attribués par chacun d'eux, ce qui a pour effet de réduire les variations entre correcteurs (Blais & Auger, 1994). Ces stratégies ne sont toutefois rentables que pour des épreuves à grande échelle qui comptent sur les services d'importantes équipes de correcteurs (Braun, 1988).

Les procédures indirectes, de leur côté, peuvent offrir des garanties de fidélité plus élevées, un questionnement ciblé sur des contenus spécifiques, le recours à différentes situations contextuelles, une «couverture» notionnelle relativement étendue, de même qu'une correction en général plutôt simple et peu coûteuse.

Cependant, on reproche à cette forme d'épreuve de laisser croire aux candidats que la maîtrise de l'écriture est moins importante que la maîtrise des sous-habiletés fragmentées qui sont testées. Cette question de l'impact de la présence d'une mesure atomisée sur les perceptions des candidats — et parfois même sur l'enseignement — est jumelée à une question théorique qui est surtout traitée dans la recherche sur les langues secondes: la compétence

langagière est-elle constituée d'un facteur général, unique et indivisible, ou d'un ensemble de sous-habilités qui peuvent être enseignées et mesurées séparément? En d'autres termes, le tout est-il plus grand que la somme des parties prises isolément? (Voir Oller, 1978, 1983.)

Les procédures indirectes sont par ailleurs accusées, par les linguistes et les enseignants des langues surtout, de viser un aspect réduit de la compétence à écrire (uniquement la maîtrise du code) et de présenter une image déformée des compétences du candidat (voir Benton & Kiewra, 1986). Comme ces procédures utilisent traditionnellement des items qui portent sur l'usage linguistique uniquement (l'orthographe, le vocabulaire, la syntaxe, par exemple), il n'est pas étonnant qu'on hésite à accepter qu'elles permettent de mesurer la totalité de la compétence à écrire. Elles ne permettent des interprétations qu'à propos de ce qui est inclus dans l'instrument utilisé.

### **L'objectif de la recherche**

On accepte généralement que les items à réponse choisie permettent de formuler des interprétations valides à propos de la maîtrise des aspects normatifs de la langue (orthographe, ponctuation, lexicale, syntaxe). Qu'en est-il toutefois lorsqu'on désire mesurer, à l'aide de telles procédures (indirectes), les autres aspects de la compétence à écrire: les habiletés dites «supérieures», telle la maîtrise de la cohésion sémantique, par exemple? Tant que de tels projets n'ont pas été tentés, on ne peut apporter de réponse à cette question. C'est ce que propose la présente recherche, de façon exploratoire.

Dans une étude sur la mesure de l'habileté à écrire, Benton et Kiewra (1986) ont interprété les résultats d'items nécessitant des habiletés d'organisation. En ajoutant, à des items portant sur le code, des items mesurant des habiletés d'organisation (résolution d'anagrammes, réarrangement de phrases), ils ont obtenu une corrélation plus élevée avec des procédures directes de mesure qu'en employant les items sur le code uniquement. Ils terminent leur étude en signalant que des items peuvent être conçus pour mesurer les habiletés organisationnelles et que l'élaboration d'items qui «captent» ces aspects de la compétence à l'écrit peut s'avérer très utile.

Ackerman et Smith (1988) ont comparé des résultats de procédures directes et indirectes de mesure à la lumière du modèle cognitif de Hayes et Flower (1980), qui définit, pour la rédaction, des processus

d'organisation, de transformation et de révision. Ils montrent que des tâches différentes nécessitent des traitements cognitifs différents et indiquent que la différence des résultats obtenus par les procédures directes et indirectes, et surtout par des procédures indirectes différentes entre elles (items à réponse choisie et items à réponse construite), dépend non seulement du contenu de la tâche, mais aussi de sa forme. Ils concluent que les méthodes indirectes peuvent être modifiées pour mesurer les aspects de la compétence à l'écrit qui requièrent des habiletés organisationnelles supérieures.

À la lumière des travaux de Benton et Kiewra (1986), dont il vient d'être question, il apparaît que les stratégies de cohésion sémantique, pourraient être mesurées à partir de procédures indirectes. Cette information, de même que les considérations sur les processus cognitifs et les procédures indirectes de mesure formulées par Ackerman et Smith (1988), ont conduit à la rédaction de prototypes d'items à réponse choisie. Dans la présente recherche, ces prototypes d'items sont mis à l'essai dans le but d'explorer la validité des interprétations que les scores obtenus permettent de faire à propos d'un aspect spécifique de la compétence à écrire, les macrostratégies de cohésion sémantique.

La validité des inférences faites à partir de données est une problématique centrale de la méthode scientifique en général et elle se révèle cruciale dans le domaine de l'éducation. La nature des conséquences, pour la formation des individus ainsi que pour la qualité des programmes, des décisions qui se fondent sur des «mesures» explique qu'on s'intéresse à la qualité de l'information et à la justesse de son interprétation. À l'heure où l'enseignement de l'écriture devient une préoccupation grandissante, tant au Québec qu'ailleurs, la recherche portant sur la validation de moyens efficaces et économiques de mesurer et d'évaluer la compétence à écrire revêt toute son importance.

## Le cadre conceptuel

S'il est vrai que la validité des procédures actuelles de mesure de la compétence à écrire pose de nombreux problèmes, c'est parce que les liens entre les procédures directes ou indirectes et la conceptualisation de cette compétence n'arrivent pas à s'établir clairement. En effet, il apparaît que l'ensemble des interactions entre les multiples aspects de la compétence à écrire et les divers contextes de mesure forme un champ d'investigation vaste à propos duquel les connaissances actuelles sont partielles. Avant de

proposer une démarche qui amènerait des éléments de réponse au problème abordé, il est nécessaire d'explicitier plus à fond les concepts qui encadreront et guideront l'interprétation des résultats. Dans un premier temps, nous allons situer le concept de «compétence à écrire» et, dans un deuxième temps, nous allons aborder les éléments qui concourent à assurer la validité des interprétations réalisées. Ces deux volets conceptuels serviront à interpréter non seulement les convergences entre deux procédures de mesure, mais également les divergences.

### **La compétence à écrire**

Le concept de compétence à écrire est utilisé dans différents domaines, tels que l'éducation, la psychologie et la linguistique, et on rencontre différentes façons de l'aborder. En fait, on parle de plus en plus d'aspects de la compétence à écrire plutôt que de compétence comme telle (Lefebvre & Turcotte, 1989). De plus, chacun de ces aspects peut se décomposer en une multitude de facettes qui sont autant d'éléments différents de la compétence à écrire. On devrait donc, en premier lieu, reconnaître que la compétence à écrire – et c'est ce qui la rend difficile à mesurer en soi – est multidimensionnelle.

Une première façon de concevoir la compétence à écrire est de la distinguer de l'incompétence. Cette approche, qui implique l'établissement d'un critère de distinction, et donc d'une «opérationnalisation» de ce que constitue la compétence, est propre à la mesure permettant une interprétation critériée. Définir la compétence à écrire en termes «opérationnistes», c'est-à-dire en rapport avec la performance à atteindre, pose toutefois des problèmes de validité. Cronbach (1988) le résume bien en montrant que selon cette conception, l'interprétation commence et se termine avec une définition très spécifique des sortes de tâches que le candidat devrait affronter; un test qui constitue un échantillon du domaine défini permet alors, théoriquement, de formuler des inférences solides. Il poursuit en affirmant que cette perspective est cohérente, mais imprévoyante: pour comprendre la faiblesse d'une performance, à des fins de remédiation, pour améliorer les méthodes d'enseignement et pour forger des domaines plus fonctionnels, la conceptualisation d'une compétence autrement que par la simple performance est nécessaire.

Une autre approche consiste, au contraire, à distinguer la compétence de la performance. Ainsi, Chomsky (1965) définit la compétence langagière comme la connaissance qu'un individu a de la langue et la performance comme l'emploi de la langue dans des situations concrètes.

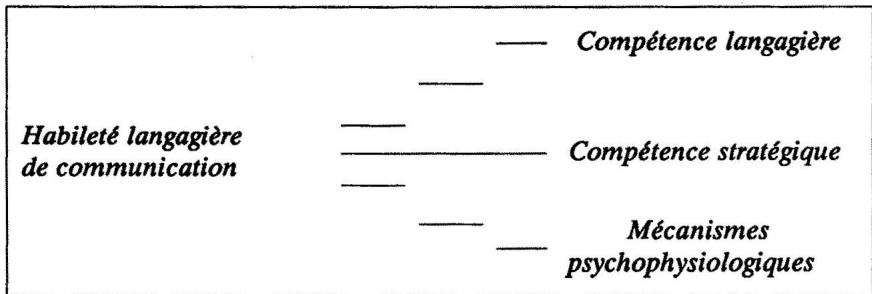
Lebrun (1987), dans ses travaux sur la compréhension en lecture, analyse le concept de «compétence» en retenant que le sociolinguiste Hymes (1974) «émet son scepticisme quant à la nécessité de conserver, dans une discipline qui touche les sciences de l'homme, la distinction entre compétence et performance. Il ne faut récupérer que le terme de "compétence" et en généraliser la définition, la rendant dépendante à la fois du savoir tacite et de l'habileté dans l'utilisation.» Il ne s'agit alors plus d'identifier de façon «opérationniste» la compétence à des critères de performance, mais plutôt de recourir à une conception qui tient compte à la fois de la dimension virtuelle et de la dimension contextualisée de la compétence.

Une autre façon de concevoir la compétence à écrire consiste à proposer une définition de la compétence à écrire à partir d'un facteur général unique, qui repose sur le postulat que les interactions entre les diverses sous-habilités sont aussi importantes que ces sous-habilités elles-mêmes et qu'en tentant de les isoler, on s'éloigne de l'habileté réelle (Oller, 1978, 1983). Cela pourrait raisonnablement bien décrire la compétence à écrire. En effet, on imagine mal l'écriture comme une activité qui recourt à des éléments cloisonnés, sans liens les uns avec les autres. Bien que cette approche pourrait constituer une objection sérieuse à certaines utilisations de la mesure indirecte, qui procède en isolant les sous-habilités, elle n'écarte toutefois pas le besoin d'inclure à une définition de la compétence à écrire une description de l'ensemble des sous-habilités qui la composent. Ce n'est pas parce que ces sous-habilités ne fonctionnent pas isolément qu'elles n'existent pas; avant de décrire le tout comme un ensemble d'interactions entre ses éléments constitutants, il apparaît nécessaire d'identifier et de décrire ces éléments constitutants. Or, cette étape préliminaire fait encore l'objet des recherches actuelles sur l'écriture.

Bachman (1990) propose un modèle langagier, qu'il nomme «habileté langagière de communication» (*communicative language ability*), adapté de Canale et Swain (1980) et de Canale (1983), qui inclut trois composantes: la compétence langagière, la compétence stratégique, les mécanismes psychophysiologiques (voir figure 1). La compétence langagière se définit comme un ensemble de savoirs spécifiques qui sont utilisés en communication par le langage. La compétence stratégique est la composante la plus originale de ce modèle en ce qu'elle permet d'expliquer la distinction entre la compétence langagière et l'utilisation de cette compétence. Elle est définie comme la capacité mentale de mettre à exécution la compétence langagière dans l'utilisation communicative contextuelle du langage. Enfin,



les mécanismes psychophysiologiques se réfèrent aux processus neurologiques et psychologiques qui participent à l'exécution du langage comme phénomène physique audible (voix) ou visible (texte). Les trois compétences du modèle sont chacune à leur tour segmentées en un certain nombre de sous-composantes. Dans la présente étude, ce modèle général a été adapté afin de décrire le concept de compétence à écrire.



**Figure 1 - L'habileté langagière de communication selon Bachman (1990)**

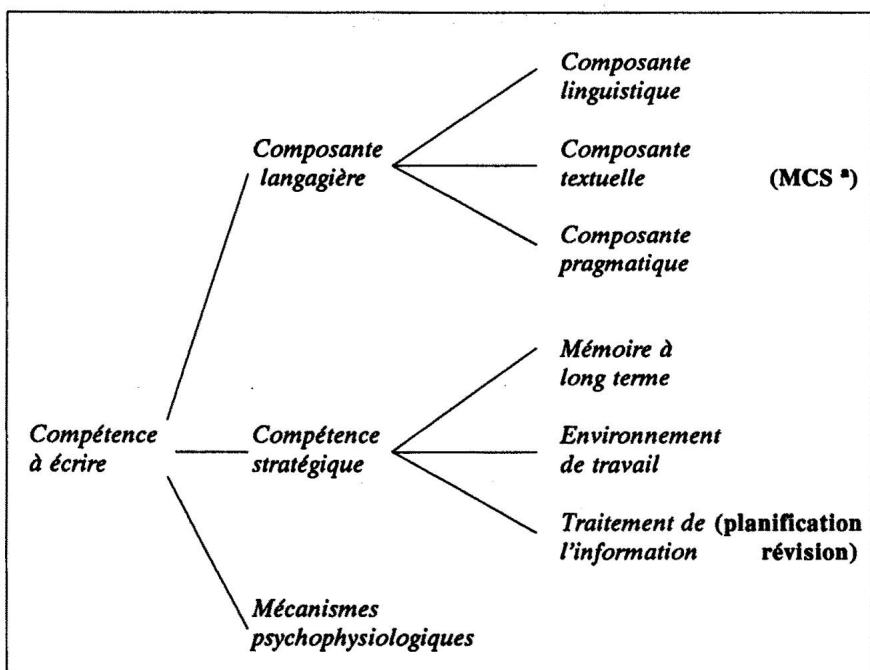
La représentation que nous utilisons (voir figure 2) reprend donc chacune des trois compétences proposées par Bachman (1990). Nous décrivons toutefois les composantes de la compétence langagière selon une organisation inspirée de Lefebvre & Turcotte (1989) et de Lebrun (1987). On y retrouve une composante linguistique (les différents usages, tels le vocabulaire, la morphologie, la syntaxe et l'orthographe), une composante textuelle (la cohésion sémantique et l'organisation rhétorique) et une composante pragmatique (les fonctions discursives et l'adéquation au contexte).

Pour la compétence stratégique, les composantes de notre modèle répondent à ce que Bachman (1990) nomme une définition psycholinguistique. Ces composantes sont issues notamment des travaux de Hayes et Flower (1980), dont la modélisation cognitive des processus d'écriture a été reprise depuis dans diverses recherches (voir, par exemple, Bereiter & Scardamalia (1987) pour une étude des aspects didactiques, ou Ackerman & Smith (1988) pour une application en évaluation des apprentissages). Trois espaces y sont décrits: la mémoire à long terme du scripteur (connaissance du sujet, des destinataires éventuels et de certains schémas textuels), l'environnement de travail (la représentation de la tâche d'écriture et le texte déjà produit) et enfin, le coeur du modèle, le traitement



de l'information par le scripteur (génération d'idées, planification, organisation, transcription et révision).

En ce qui a trait aux mécanismes psychophysiologiques, notre modèle ne tente pas d'en définir les composantes, car ces aspects de la compétence à écrire ne sont pas abordés à l'intérieur de la présente étude. Ces mécanismes sont néanmoins nécessaires au concept général et se rapportent aux capacités perceptuelles (la vue) et motrices (l'écriture manuscrite ou dactylographique) sollicitées chez le scripteur.



a. MCS : macrostratégies de cohésion sémantique

**Figure 2 - Composantes de la compétence à écrire**

On retrouve entre parenthèses, à la figure 2, les éléments à partir desquels on a effectué la spécification des items sur lesquels porte notre étude (voir la méthode): les macrostratégies de cohésion sémantique (MCS), le processus de révision et le processus de planification.

Les macrostratégies de cohésion sémantique font partie de la composante textuelle de la compétence langagière. Il s'agit du troisième niveau

d'un modèle portant sur la cohésion sémantique proposé par van Dijk et Kintsch (1983). Ce modèle a été conçu au départ pour la lecture, mais ses auteurs ont soutenu que les structures requises à la lecture sont les mêmes à l'écrit. On y distingue les stratégies propositionnelles (qui fonctionnent sur le plan des mots et des phrases), les stratégies de cohérence locale (qui relient des phrases successives) et les macrostratégies (qui agissent sur le plan des paragraphes ou du texte entier).

L'analyse du processus de révision effectuée par Nold (1981) montre que ce processus intervient tout au long de la démarche d'écriture (et non uniquement à la fin), que la difficulté de la tâche modifie les stratégies de révision employées, que l'information utilisée par le rédacteur lors de la révision ne se trouve pas uniquement dans le texte et, enfin, que les révisions apportées par le scripteur n'ont pas nécessairement pour effet d'améliorer le texte.

Quant au processus de planification, Flower et Hayes (1981) montrent qu'il n'est pas constitué d'une succession linéaire de plans, mais plutôt d'un ensemble d'interactions entre plans. Il est intéressant de retenir que les objets qui nécessitent des décisions de planification de la part du scripteur se rapportent autant au texte lui-même qu'à des opérations sur ce dernier, telle celle qui consiste à en assurer la cohésion sémantique (organisation, place des éléments, orientations, connexions).

### **Les supports à la validité des interprétations**

Les trois types de validité qui sont traditionnellement retenus, lors de la mise au point d'instruments de mesure en éducation et en psychologie, sont la validité de contenu, la validité en référence à un critère et la validité de construit (APA, 1966, 1974). Depuis peu, cependant, on traite davantage la validité comme un concept unitaire se référant à la pertinence, à la signification et à l'utilité des inférences faites à partir des scores à un test. Il est alors moins question de types de validité que de types d'évidences qui supportent la validité (APA, 1985). Plusieurs, comme Messick (1989), parlent alors de validation, qui est l'action d'accumuler ces évidences. Cet auteur propose un processus général de validation dont les composantes sont la représentativité du contenu, la consistance structurelle, la validité nomologique, la validité en référence à un critère et la portée nomothétique. Chacune de ces composantes est décrite brièvement et l'une d'entre elles sera reprise au moment d'interpréter les résultats de l'étude.

La représentativité du contenu est assurée en soumettant les items à des experts du contenu qui évaluent leur relation au domaine conceptuel à

propos duquel on tente d'établir des inférences. Traditionnellement, on a abordé cette composante de la validation en employant l'expression «validité de contenu». Depuis Cronbach (1971), on préfère parler de représentativité de contenu, car la validité des inférences doit être démontrée à partir des scores recueillis par un instrument de mesure et non à partir de sa démarche de conception (Angoff, 1988).

La consistance structurelle porte sur la relation entre le tout et ses parties. Dans le contexte qui nous intéresse, on pourrait étudier, par exemple, la relation qui existe entre un score attribué à l'ensemble du texte et des scores attribués aux différents aspects de ce dernier.

La validité nomologique correspond à ce qu'on nomme habituellement la validité de construit, c'est-à-dire les liens entre les scores à un test et le construit cognitif qui leur correspond. Lorsque des procédures de mesure différentes fournissent avec constance des résultats semblables, on peut croire, en toute légitimité, qu'elles mesurent «la même chose». Dans le même sens, peut-on dire que les procédures directes et indirectes mesurent le même «trait»?

La validité en référence à un critère pose la question de l'établissement de «critères» de comparaison pertinents. Par exemple, les notes obtenues aux cours de langue maternelle sont souvent considérées comme des critères peu valables parce qu'elles sont, la plupart du temps, constituées à partir de mesures effectuées en classe par les enseignants, mesures qui, d'une part, ne concernent pas uniquement l'écriture et dont la validité, d'autre part, n'est pas démontrée (Charney, 1984).

La portée nomothétique d'un test nous amène à examiner comment le score obtenu peut dépendre de facteurs externes, qui devraient théoriquement être étrangers à la mesure. En trouvant des moyens de réduire l'influence des différents «parasites» connus, les chercheurs pourraient accroître la validité globale des mesures de la compétence à écrire. En ce qui a trait aux procédures directes de mesure, Miller et Crocker (1990) soulignent que malgré plusieurs efforts, peu de progrès ont été enregistrés à ce chapitre. De plus, selon Messick (1989), la portée nomothétique est constituée par l'accumulation de résultats convergents entre différentes recherches de validation, c.-à-d. d'abord entre mesures différentes d'un même construit et entre mesures d'un même construit dans des conditions différentes, et ensuite entre des mesures de construits différents mais théoriquement reliés.

Finalement, il faut ajouter que même en assurant à une mesure toutes les qualités dont il a été question, cela ne permet pas d'affirmer que n'importe quelle utilisation de cette mesure est valide et peut permettre n'importe quel type d'évaluation. Messick (1989) insiste donc sur le fait qu'une démarche complète de validation doit aussi porter sur les conséquences de l'évaluation.

Ainsi, différents types d'évidences peuvent théoriquement supporter la pertinence, la signification et l'utilité d'interprétations faites à partir de scores. On ne saurait cependant, dans le cadre de la présente étude, traiter de manière équivalente tous les types d'évidences mentionnés, chacun d'eux requérant respectivement une ou plusieurs démarches d'investigation, ce qui pourrait constituer un programme de recherche en entier. Cronbach (1988) rappelle qu'il s'agit d'une tâche qui, en fait, n'est jamais terminée. La composante du processus de validation à laquelle cette étude s'intéresse particulièrement est la validité nomologique (ou de construit). C'est donc dire qu'on tente surtout d'établir des liens entre les scores aux items mis à l'essai et la représentation du concept de compétence à écrire.

## La méthode

Les études de validation des procédures directes et indirectes de mesure reposent habituellement sur une démarche quantitative d'analyse corrélationnelle. Dans la présente recherche, nous nous éloignons de la démarche «traditionnelle». La démarche choisie consiste à examiner, de manière qualitative, les données empiriques à la lumière des concepts cognitifs et linguistiques pertinents. Ainsi, la démarche méthodologique privilégie l'interprétation de scores individuels à partir de cas exemplifiés, plutôt que les tendances se dégageant des relations entre les scores issus de deux situations de mesure. Comme la présente recherche se situe au tout début de la démarche de validation des items, il est justifié de commencer par observer quels résultats sont obtenus avec quelques individus avant de penser à une étude auprès d'un grand groupe.

Cette approche n'est toutefois pas uniquement motivée par une économie du nombre de sujets. Les coefficients de corrélation «de faibles à modérés» observés lors des études de validation de procédures de mesures directes et indirectes de la compétence à écrire ont souvent mené des chercheurs à des conclusions mitigées quant au construit mesuré (ou aux construits mesurés). Une approche qualitative peut permettre de mieux décrire la complexité des phénomènes en jeu. Dans la présente recherche,

des items portant sur les macrostratégies de cohésion sémantique (MCS) sont mis à l'essai auprès d'un groupe d'élèves du collégial. Les interprétations que permettent les réponses à ces items sont ensuite analysées en regard d'autres données: l'analyse de textes produits par ces élèves, de même que leurs réponses à un questionnaire d'opinion. On peut ainsi voir s'il existe une convergence dans les données, mais aussi tenter d'interpréter, à l'aide du cadre conceptuel, ce qui se passe effectivement lorsque les données ne convergent pas.

Les items mis à l'essai se regroupent en trois séries. Ils nécessitent tous l'utilisation par le candidat de macrostratégies de cohésion sémantique (MCS) sur le plan du paragraphe. De plus, chaque série d'items tente d'aborder la mesure des MCS selon l'un ou l'autre des deux processus cognitifs suivants: la planification ou la révision. Ces processus ne sont pas directement visés par la mesure; ils fournissent plutôt aux items des contextes de mesure différents. On tente ainsi d'adapter les trois formats d'items à des situations d'écriture variées.

La première série comporte cinq items, qui sollicite les MCS en plaçant le candidat en situation de planification. Les tâches consistent à ordonner les idées d'un brouillon afin de former un paragraphe qui assure une progression sémantique adéquate. On demande au candidat de choisir parmi quatre choix d'ordonnements possibles. Il s'agit donc d'items à choix multiples conventionnels.

#### Série 1 (item 1)

À chacun des numéros suivants, on te propose une liste d'idées qui sont tirées d'un brouillon. Tu dois former un paragraphe logique et cohérent à partir de ces idées en les plaçant en ordre. Encerle, parmi les choix de réponses proposées, le chiffre qui correspond à la suite d'idées la plus appropriée.

1. A) Le festival folklorique est tenu la première semaine d'août.  
B) Tous les habitants portent le costume traditionnel.  
C) L'an dernier, malheureusement, le mauvais temps a nui au succès de la fête.  
D) Chanteurs, danseurs et amuseurs publics se succèdent sans arrêt.

- 1) A - D - B - C
- 2) C - A - D - B
- 3) B - C - A - D
- 4) D - B - C - A

La deuxième série comporte vingt-cinq items, qui situent à nouveau le candidat dans un contexte de planification. Les tâches consistent à discriminer diverses idées se rapportant à un sujet général et à découvrir le thème qui s'en dégage, afin de sélectionner uniquement les idées qui permettent de rédiger un paragraphe cohérent sur le plan sémantique. Pour chaque idée présentée, le candidat doit indiquer si elle doit être retenue ou non. Plusieurs items sont donc liés à un sujet. Ce format s'apparente aux items à multiples vrais-faux décrits par Haladyna (1992).

**Série 2 (items 1 à 4)**

Dans chacun des numéros suivants, on présente une liste d'idées, extraites du brouillon d'un paragraphe. Afin de respecter la logique du paragraphe, il est possible que certaines idées ne puissent pas être retenues. Indique, pour chacune des idées, si elle devrait être retenue ou non afin d'écrire un paragraphe d'au moins trois (3) idées, cohérent sur le plan du contenu. Encerle en bout de ligne le chiffre qui convient.

Rappel: tu dois retenir au moins trois (3) idées par numéro.

**Idee retenue pour  
le paragraphe**

**OUI NON**

A. 1) La ville se dépeuple au profit des plages et des montagnes.	1	2
2) C'est en juin que débutent les vacances.	1	2
3) L'hiver est parfois très rude.	1	2
4) La ruée de vacanciers rend les routes très achalandées.	1	2

La troisième et dernière série d'items en compte neuf. Le candidat y est placé en situation de révision. Il doit utiliser les MCS afin de réaliser le «découpage» en paragraphes d'un texte, en assurant la cohésion sémantique interne des paragraphes et la cohésion des paragraphes entre eux. Les neuf items sont liés au même texte initial. Il s'agit à nouveau d'un format d'item de type multiples vrais-faux.

Après leur rédaction, les items ont été révisés par deux experts linguistes afin d'en assurer la pertinence en regard des construits théoriques visés (les MCS, d'une part, et les processus de planification et de révision, d'autre part). Cette révision n'avait cependant pas pour but de fixer un seuil de réussite, ni même d'établir un jugement quant à la difficulté des items.

## Série 3 (items 1 à 9)

Le texte qui suit a été écrit d'un premier jet. Tu dois le découper en paragraphes logiques et cohérents. Pour chacune des lettres entre crochets, indique en bout de ligne si cet endroit devrait correspondre à une coupure. Encerle le chiffre qui convient.

## COUPURE EN PARAGRAPHE

OUI NON

Tant qu'on ne l'a pas visité, on ne connaît souvent d'un pays que le nom de ses villes importantes. [A] Au-delà des zones urbaines, toutefois, se cache souvent l'âme la plus intacte d'une région, la moins transformée par l'industrialisation de notre époque. [B] Chez moi, dès qu'on s'éloigne des grands centres, les chemins bordés d'arbres soulignent les particularités du relief, longeant les cours d'eau, épousant les courbes du paysage. [C] Le spectacle est si agréable qu'on comprend ceux qui, il y a très longtemps, ont choisi cette contrée pour s'y fixer. [D] Et lorsque de loin on aperçoit les pierres des villages, on se croirait revenu des siècles en arrière. [E] Cependant, une antenne de télévision ou un garage rappellent brusquement les rêveurs à la réalité. [F] La lutte que se livrent les âges prend parfois des formes plus brutales: une usine, un cimetière d'autos ou une centrale électrique apparaît dans un détour. [G] La note pittoresque, malgré tout, domine dans le souvenir qu'on ramène. [H] Si vous passez par ici un jour, allez bien sûr vous promener dans les grandes villes. [I] N'oubliez pas, cependant, d'en sortir: celui qui cherche à découvrir les racines d'un peuple, ses traditions et son histoire, gagne à parcourir les campagnes et les villages.	[A]	1	2
	[B]	1	2
	[C]	1	2
	[D]	1	2
	[E]	1	2
	[F]	1	2
	[G]	1	2
	[H]	1	2
	[I]	1	2

La mise à l'essai des items devait avoir lieu auprès de 21 élèves de première année du collégial du collège de l'Assomption, 14 filles et 7 garçons. Ce collège est un établissement d'enseignement privé situé en banlieue de Montréal, qui offre des programmes d'enseignement général et

professionnel. Sa clientèle, relativement homogène, provient surtout de familles à revenu moyen et de langue maternelle française. Le choix pour la mise à l'essai d'élèves du collégial, c'est-à-dire d'individus atteignant l'âge adulte, permet de traiter la compétence à écrire indépendamment du développement de l'individu, aspect dont on doit tenir compte lorsqu'on fait porter les observations sur des individus plus jeunes. Il ne s'agit pas de prétendre que «l'individu collégial» ne se développe plus, loin de là, mais plutôt qu'il a atteint une maturité psychophysiologique suffisante pour qu'il soit injustifié d'expliquer une lacune observée sur le plan de la compétence à écrire par une référence au stade de développement, alors que cette hypothèse serait valable s'il s'agissait d'une étude portant sur la compétence à écrire de jeunes enfants. Pour ce qui est du choix de l'établissement d'enseignement, il est motivé par les travaux antérieurs qui y ont été menés et qui l'ont doté d'un instrument de mesure de la compétence à écrire, administré en début d'études collégiales (Fagnant et Rousseau, 1990). Dans la présente recherche, ce sont les textes produits lors de la partie rédaction de ce test qui sont recueillis et analysés.

Le professeur du groupe d'élèves participant à la mise à l'essai rapporte que ce groupe ne diffère pas des trois autres groupes auxquels il donne le même cours, ni des groupes qu'il a rencontrés dans son enseignement des dernières années. Il précise qu'il y a des élèves de la classe qui, comme dans la plupart des groupes, se démarquent par la qualité de leurs rédactions, et d'autres dont la qualité des rédactions est généralement faible. On peut donc croire que la compétence à écrire des élèves de ce groupe est relativement diversifiée.

La mise à l'essai des items a été planifiée conjointement avec le professeur du groupe d'élèves et comprend trois étapes intégrées à l'enseignement, chacune se déroulant à l'intérieur de périodes de cours, espacées entre elles d'une semaine: présentation, mise à l'essai et retour.

Lors de la présentation, on a rencontré les élèves dans les vingt premières minutes du cours et on leur a expliqué sommairement la démarche de recherche mise en place. On a également insisté sur l'aspect formateur que pouvait représenter pour eux cette participation, d'autant plus que la mise à l'essai est suivie d'une période de discussion portant sur la compétence à écrire à la lumière des résultats obtenus. Cette sensibilisation à la démarche et l'accent mis sur son caractère formatif avaient pour but de susciter l'intérêt des élèves et d'éviter ainsi qu'ils ne donnent pas leur plein effort lors de la mise à l'essai des items.



La mise à l'essai a eu lieu une semaine après la présentation et a duré toute la période de cours, c'est-à-dire 55 minutes. Les élèves ont été invités à répondre sur les cahiers qui leur étaient distribués, sans consulter quelque matériel que ce soit. Il a été expliqué oralement, au début de la période, qu'on ne répondrait à aucune question de clarification portant sur les items. Les élèves devaient indiquer sur les questionnaires mêmes ce qui pouvait poser problème. Il leur était précisé qu'après avoir terminé de répondre à tous les items et aux questions qui suivaient, ils pouvaient remettre leur cahier et quitter le local.

Le retour, qui a eu lieu une semaine après la mise à l'essai, a duré lui aussi 55 minutes (une période de cours) et a pris la forme d'une discussion où les items étaient tour à tour présentés à l'aide d'un rétroprojecteur. En présentant les réponses attendues et celles qui ont été observées, on a sollicité les commentaires et les questions des élèves. Un enregistrement sonore de cette rencontre a été réalisé, afin de recueillir les commentaires des élèves.

Le collègue a mis à notre disposition les textes produits par les élèves du groupe lors du test d'entrée administré au début de la session d'automne, soit environ six mois avant la mise à l'essai des items. Ce délai entre la production des textes et la mise à l'essai n'est certes pas une condition souhaitable, mais il serait étonnant que l'aspect de la compétence à écrire visé par les items, soit les macrostratégies de cohésion sémantique, ait tellement changé durant cette période.

Par ailleurs, l'encadrement de la séance d'administration du test confère à ces textes certaines qualités qu'il aurait été difficile d'obtenir autrement, sans abuser du temps de classe offert par le professeur. En effet, on peut être certain que ces textes ont été écrits par les élèves dans des conditions uniformes et connues, qu'ils ont été réalisés individuellement et non avec l'aide d'un tiers, sans aucune documentation, et qu'il s'agit de productions pour lesquelles un effort sérieux a été fourni.

Un bref questionnaire a été joint en dernière page du cahier de l'élève, afin de recueillir une information écrite de la part des élèves sur le déroulement de la mise à l'essai et sur les items. Les trois questions posées étaient les suivantes:

1. *As-tu déjà répondu à des questions à choix de réponses de ce genre? Si oui, dans quel contexte?*

2. *Crois-tu que les questions auxquelles tu viens de répondre peuvent fournir une information valable sur ton habileté à écrire? Explique brièvement.*
3. *Comment qualifierais-tu l'effort que tu as fourni en répondant à ce test qui, rappelons-le, n'influçait pas ta note de ce cours de français?*

Le premier type de données recueillies est constitué des réponses aux items. Pour chaque élève, on a calculé un score total pour les macrostratégies de cohésion sémantique, en accordant un point par réponse correcte. En ce qui concerne le second type de données, les textes d'élèves, il a fait l'objet d'une retranscription dactylographiée pour en faciliter l'analyse. Cette retranscription n'a changé ni les mots ni les formulations, mais a corrigé, dans la plupart des cas, les erreurs orthographiques afin qu'elles n'entravent pas la lecture. Les commentaires qui ont été recueillis à l'aide du questionnaire et de l'enregistrement, soit le troisième type de données, ont fait l'objet d'une analyse de contenu et les idées qui s'en dégagent ont été regroupées par thèmes.

## Les résultats

### Les commentaires des élèves

Les commentaires des élèves ont été recueillis grâce aux questions incluses dans le cahier de l'élève et à la discussion lors du retour ayant suivi la mise à l'essai. Ces données sont pertinentes à l'interprétation des scores, car elles auraient pu révéler la présence de biais importants (le manque d'intérêt ou de motivation, par exemple).

Tous les élèves du groupe affirment avoir fourni un effort valable ou normal lors de la mise à l'essai. Il apparaît donc qu'ils ont fait l'exercice avec sérieux. Sauf exceptions, les élèves ont trouvé nouvelle cette forme de questionnement. Quelques-uns rappellent qu'ils ont déjà répondu à des items à réponse choisie portant sur le français écrit, notamment lors du test d'entrée du Collège de l'Assomption (ces items portaient toutefois sur la maîtrise du code). Une élève se souvient avoir fait, au secondaire, des exercices de remise en ordre de bouts de texte, mais elle ne précise pas s'il s'agissait d'items à réponse choisie.

Les élèves croient, pour la plupart, que les items auxquels ils ont répondu peuvent fournir des renseignements valables à propos d'éléments particuliers de leur compétence à écrire. Certains émettent toutefois des

réerves, précisant qu'il s'agit d'habiletés très spécifiques qui ne reflètent sans doute pas toutes les habiletés d'écriture. Lors de la mise à l'essai, aucune question de clarification ne fut posée par les élèves au sujet des tâches de mesure proposées dans le cahier de l'élève. L'échange qui a eu lieu au moment du retour a permis de confirmer que les élèves comprenaient clairement les tâches qu'ils devaient accomplir.

### Les scores aux items

La distribution des scores présentée au tableau 1 indique que les items mis à l'essai ont été bien réussis par plusieurs élèves, la majorité des élèves ayant répondu avec succès à plus de 80% des items. Malgré cela, on observe quelques scores nettement inférieurs. Il est très difficile d'interpréter ces scores en eux-mêmes, autrement qu'en situant le score d'un individu par rapport au reste du groupe. Un score de 30 sur 39 pourrait sembler acceptable, mais s'avère un des plus faibles de la classe. Que peut-on en déduire ? Cet élève semble moins bien maîtriser les macrostratégies de cohésion sémantique que la plupart des autres élèves, mais comment décrire autrement sa compétence ? Si celle des autres élèves de la classe est très grande, il se peut que la sienne soit relativement grande aussi. Il se peut au contraire que les items soient trop faciles et que le groupe maîtrise en réalité plus ou moins l'objet mesuré; un score de 30 pourrait alors indiquer une lacune assez importante concernant cet aspect de la compétence à écrire. D'autres données sont nécessaires afin d'interpréter ces scores.

**Tableau 1**

**Distribution des scores pour le total  
des trois séries d'items (39 items, 21 sujets)**

Score	23	25	27	28	29	30	31	32	33	34	35	36	38
Fréquence	1	1	1	1	1	1	1	1	2	2	6	1	2

### Les textes des élèves

L'analyse des textes de chacun des 21 élèves du groupe sur le plan des macrostratégies de cohésion sémantique (MCS) a été effectuée en comparant le score obtenu à ce qui peut être perçu dans le texte, afin de

tenter d'en dégager des interprétations. Les textes ont ainsi été regroupés en quatre catégories: (a) avec MCS efficaces, (b) avec MCS efficaces et MCS inefficaces, (c) avec MCS inefficaces, (d) sans MCS observables.

**Tableau 2**

**Distribution des textes d'élèves en catégories**

Catégorie	- a - avec MCS* efficaces	-b- avec MCS efficaces et MCS inefficaces	- c - avec MCS inefficaces	- d - avec MCS observables
	36	35	33	35
	34	35	30	23
	38	35	29	32
	25	35	28	31
	38	34		27
				33
				35
<b>n=</b>	<b>5</b>	<b>5</b>	<b>4</b>	<b>7</b>

\* MCS: macrostratégies de cohésion sémantique

Comme il n'est pas possible de reproduire ici l'analyse complète de chacun des 21 textes, nous avons choisi deux textes représentant chacune des catégories afin d'illustrer la démarche (cependant, les analyses de tous les textes ont contribué à la discussion qui suit la présentation des résultats). Les textes présentés ci-après sont examinés à partir d'extraits jugés révélateurs de leurs principales caractéristiques sur le plan des MCS.

*a) Textes avec MCS efficaces*

Le premier extrait présenté est celui d'une des deux élèves ayant obtenu le score le plus élevé du groupe, 38 points sur un maximum possible de 39 (élève 6). Son texte est sans lacune sur le plan de la cohésion sémantique d'ensemble.

En un troisième temps, plusieurs ont la sage idée de tenir compte des besoins du marché du travail. Dans ce cas, il devient plus facile pour certains de se trouver un emploi rapidement. D'un autre côté, il ne faut pas

seulement se fier aux disponibilités d'emplois, car le jeune travailleur n'aimera peut-être pas ce qu'il devra effectuer ou il ne sera peut-être pas très bon dans cette carrière. Ces deux derniers facteurs affecteront son rendement et il ne pourra pas ou il n'aura pas le goût d'avancer plus loin dans ce domaine.

Pour conclure, je crois donc que chaque étudiant devrait faire son choix en se basant autant sur ses goûts, que sur ses aptitudes et que sur les besoins du marché du travail. En tenant compte de ces trois facteurs à part égale, le jeune pourra avoir une carrière qu'il aime, dans laquelle il est bon et où il a beaucoup de chances d'obtenir un emploi intéressant. (Élève 6, paragraphes 4 et 5.)

Le paragraphe d'introduction de ce texte annonce avec clarté et précision les trois idées qui sont développées dans le corps du texte. Les idées sont reprises en conclusion, qui est une véritable synthèse et non uniquement une répétition, comme le montre l'extrait reproduit. La structure du paragraphe 4 est plus complexe que les structures observées dans la plupart des textes du groupe, en ce sens qu'elle dépasse la simple exposition de l'idée et intègre, sans rupture de cohésion, le développement d'un contre-argument. Le paragraphe 2 présente les mêmes qualités. Cette maîtrise des macrostratégies de cohésion sémantique, même complexes, s'accorde assez bien avec l'interprétation qu'on peut faire du score presque parfait aux items.

L'élève suivant (élève 10) a obtenu un score relativement faible (par rapport à l'ensemble), soit 25. Cette information est plutôt divergente de celle qui se dégage de la lecture de son texte. En effet, ce dernier affiche une assez bonne maîtrise de certaines macrostratégies de cohésion sémantique simples. L'introduction qui suit permet d'observer comment ces stratégies sont mises en oeuvre.

De plus en plus, le gouvernement du Québec songe à augmenter l'âge minimal pour l'obtention du permis de conduire à 18 ans. Nous savons que l'âge légal pour obtenir ce permis est actuellement à 16 ans. Personnellement, je suis tout à fait contre cette idée. Je diviserai mon argumentation en 3 points: soit la responsabilité, la maturité et la liberté face aux parents. (Élève 10, paragraphe 1.)

Dès le départ, les idées sont clairement présentées et annoncent l'organisation du texte, à l'intérieur duquel chacune des trois idées principales énumérées fait l'objet d'un paragraphe. Les paragraphes sont simplement reliés par les marqueurs: «Pour débiter...», «Deuxièmement...», «Troisièmement...», et la conclusion est une reprise

paraphrasée de la dernière phrase de l'introduction. Ce sont donc des stratégies simples d'annonce et de reprise d'une part, et de juxtaposition d'autre part, que semble pouvoir utiliser cet élève.

*b) Textes avec MCS efficaces et MCS inefficaces*

Voici un extrait du texte d'un l'élève qui a réussi 35 des 39 items (élève 7). Il présente certaines failles sur le plan de la cohésion sémantique des structures d'ensemble, mais aussi certaines stratégies correctement employées.

Par contre les étudiants qui font leur choix de carrière selon la demande du marché du travail et qui n'aiment pas ce qu'ils ont choisi regretteront leur choix puisqu'il ne contiendra aucun de leurs goûts et aptitudes. Donc je pense que les jeunes ont davantage à choisir leur carrière en fonction de leurs goûts mais aussi en fonction des besoins du marché du travail. Les jeunes doivent tenir compte de ces deux facteurs puisque les deux réunis contribuent à la réussite de leur carrière future. (Élève 7, paragraphe 3.)

Ce dernier paragraphe du texte comprend à la fois le développement du second terme du raisonnement et la conclusion, ce qui aurait dû constituer deux paragraphes distincts pour éviter que le lecteur ne relie la conclusion uniquement à cette deuxième partie, plutôt qu'à l'ensemble du texte. De plus, le second terme du raisonnement est traité beaucoup plus brièvement que le premier, ce qui peut laisser croire que le rédacteur y accorde moins d'importance. Ce n'est peut-être pas le cas, mais comme rien ne l'indique, comment savoir? On doit toutefois souligner que la conclusion, une fois distinguée du reste, reprend bien les éléments importants du texte, éléments qui étaient annoncés en introduction: il se dégage de l'ensemble une information claire.

Le texte suivant est celui d'un des six élèves ayant obtenu le score de 35 pour les items à réponse choisie (élève 12). Il présente une situation partagée, soit des stratégies de cohésion sémantique d'ensemble correctes, mais des difficultés d'organisation interne des paragraphes, ce qui tend à indiquer que les macrostratégies de cohésion sémantique fonctionnent différemment selon le type de structure concerné (paragraphe ou texte entier).

Pour d'autres personnes par exemple, le goût et les aptitudes importent beaucoup. Si elles travaillent dans un domaine qui leur est intéressant, ils travailleront durement pour arriver à leur but: exercer la profession qu'ils aiment. Alors, leurs notes seront des meilleures et ils pourront exercer leur profession. La carrière est très importante, et personne ne devrait prendre

cela à la légère. D'après moi, la chose la plus importante est de faire quelque chose que l'on aime et dans laquelle on est bien tout en ayant les capacités nécessaires, pour effectuer la profession choisie. On ne réfléchit jamais assez quant à notre choix de carrière, car nous allons l'exercer pour la moitié de notre vie sinon plus. (Élève 12, paragraphe 3,)

Plutôt que de montrer les stratégies efficaces de structuration du texte entier présentes dans ce texte, l'extrait choisi illustre comment le rédacteur a dérogé de son propre plan en négligeant d'assurer aux paragraphes la cohésion interne nécessaire. L'introduction proposait une analyse en deux parties («Pour certaines personnes», «Tandis que pour d'autres»). Une première source possible de confusion est le non-respect de l'ordre de présentation des éléments: le premier terme de l'introduction est traité en second dans le texte, et vice versa. Le paragraphe 3, qui devrait s'en tenir au propos annoncé, comprend quant à lui (à partir de «La carrière...») divers énoncés généraux et une opinion, qui sont sans relation sémantique explicite entre eux ou avec le début du paragraphe.

*c) Textes avec MCS inefficaces*

Le texte suivant comporte certaines similitudes sur le plan des stratégies de cohésion sémantique avec celui d'autres candidats (élève 4, score de 33). En effet, c'est surtout au sujet de la signification interne des phrases et des liens entre ces dernières (stratégies propositionnelles et stratégies de cohérence locale) que les lacunes sont d'abord visibles, car elles affectent les unités de base. On se rend ensuite compte que la situation de la cohésion sémantique des structures plus globales que sont les paragraphes et le texte en entier n'est guère meilleure.

Depuis quelques années maintenant, notre entrée sur le marché du travail se fait de plus en plus complexe. Nos choix sont orientés vers des domaines qui demandent des préalables exigeants et parfois même difficiles à obtenir.

Premièrement, je crois que la meilleure solution n'est pas de s'orienter pour satisfaire les besoins du marché, car à chaque année nos besoins changent et nous devons nous accoutumer aux nouvelles situations qui se présentent à nous, alors à quoi bon de poursuivre des études qui d'un jour à l'autre pourraient nous laisser tomber!

Deuxièmement, choisir un métier, un travail ou une profession est un choix personnel. Puisque nous avons un faible pour ce travail, il nous faut l'adapter en fonction de nos goûts et aptitudes. Je ne connais personne qui travaille seulement pour les besoins du marché, les risques sont trop élevés. De toute façon à quoi bon exercer un travail qui nous est pas relié;

si de toute façon nous en serons malheureux toute notre vie. (Élève 4, paragraphes 1, 2 et 3.)

L'introduction insiste sur une idée, les préalables, qui n'est pas traitée dans le texte. La lecture de ce premier paragraphe ne contribue donc pas à la construction du sens du texte, sauf pour indiquer vaguement qu'il y est question de travail ou de formation. Quand les paragraphes subséquents débutent par «Premièrement», «Deuxièmement», «Troisièmement», rien n'indique au lecteur à quoi est reliée cette énumération. Le flou persiste lorsqu'on retrouve, comme aux paragraphes 2 et 3, une absence d'organisation à l'intérieur même des paragraphes, alors que plusieurs idées sur des sujets distincts sont juxtaposées.

Le texte suivant est celui d'un élève qui a réussi à répondre correctement à 28 des 39 items à réponse choisie, soit à plus de 70% de ces derniers (élève 16). Cela pourrait sembler acceptable, mais ce score est néanmoins l'un des moins élevés du groupe. On observe dans son texte des lacunes très marquées sur le plan de la cohésion sémantique, notamment dans les paragraphes 2 et 3.

Premièrement, je ne sais pas pourquoi l'âge pour obtenir un permis de conduire est maintenant à 16 ans. Je ne pense pas qu'un jeune de cet âge soit plus mature qu'un autre du même âge il y a trente ans. Je pense fortement que la Régie de l'Assurance Automobile du Québec devrait émettre les permis de conduire à l'âge adulte.

Ensuite, je pense qu'un jeune de 16 ans n'est pas assez mature pour avoir la responsabilité d'un véhicule, il n'a pas assez d'expérience et ne connaît pas assez le danger. Je pense par contre que la Régie de l'Assurance Automobile du Québec devrait émettre un permis temporaire d'au moins 1 an avant le vrai permis. (Élève 16, paragraphes 2 et 3.)

Dans cet extrait, on note d'abord une tentative de structuration du texte à l'aide des expressions «Premièrement» et «Ensuite», mais qui ne réussit qu'à introduire une certaine confusion. En effet, la première phrase du paragraphe 3 n'est ni plus ni moins qu'une paraphrase du début du paragraphe précédent. Ainsi, le texte ne progresse pas comme la présence du mot «Ensuite» le laissait entendre. On remarque par ailleurs que la dernière phrase du paragraphe 3, à l'aide de la locution «par contre», se pose en opposition non pas avec ce qui la précède directement, mais bien avec la fin du paragraphe précédent, faisant ainsi totalement fi de la séparation.



d) *Textes sans MCS observables*

Le score de cet élève aux items à réponse choisie (élève 1) est parmi les scores les plus élevés du groupe (35). Malgré cela, la lecture de son texte permet de constater l'absence de macrostratégies de cohésion sémantique, autant à l'intérieur des paragraphes que dans la structure d'ensemble.

Les adultes aussi sont souvent des passionnés. Dans les «arcades», la moyenne d'âge est beaucoup plus élevée que nous serions portés à penser. Comptons le nombre de pères qui s'abonnent à «Vidéoway» pour leur bon plaisir. Cependant malgré tous les inconvénients cités plus haut, je dois reconnaître que certains jeux et programmes sont enrichissants et constructifs. Mais malheureusement ils sont rares.

Pour conclure, je rappelle que je suis contre les jeux vidéo. Sans que ça paraisse, les jeux électroniques sont une des nombreuses causes de la détérioration de notre société et de l'augmentation d'échecs des ménages québécois. Là-dessus, croyez-moi je n'exagère rien. Pourquoi n'essayerions-nous pas de nous passer de ces machines qui nous font perdre toute notre créativité et beaucoup de temps ? (Élève 1, paragraphes 3 et 4.)

Le paragraphe 3 est un exemple typique d'absence de cohésion interne du paragraphe. Il renferme deux idées complètement indépendantes, et ne propose au lecteur aucune façon d'interpréter leur juxtaposition. Non seulement ce dernier risque-t-il d'être dérouté par ce brusque changement de propos, mais, ce qui porte davantage à conséquence sur le plan de la communication, il risque de l'interpréter d'une manière qui diffère totalement de l'intention du rédacteur. Le paragraphe de conclusion rappelle la prise de position initiale, mais l'appuie sur deux arguments tout à fait nouveaux, plutôt que sur la synthèse de ce qui a été développé dans le texte, et fait donc appel à un raisonnement que le lecteur ne connaît pas et qui n'est pas explicité. Cela désamorce complètement ce qui aurait pu devenir, autrement, une macrostratégie de cohésion sémantique.

L'élève suivant a obtenu le score le moins élevé du groupe (élève 2, score de 23), ce qui pourrait être une indication qu'il maîtrise moins bien les stratégies en question qu'un élève ayant obtenu un score élevé. La lecture de son texte confirme cette impression. Le troisième paragraphe est une longue digression qui comporte des contradictions internes et des contradictions avec le paragraphe de conclusion qui suit, ce qui entraîne une absence d'équilibre entre les parties du texte.

Pour vous prouver que le permis de conduire est un instrument indispensable au bien-être de l'adolescent, je vous le démontrerai avec quelques arguments très appropriés. Prenons juste comme exemple l'adolescent de 17 ans qui va au Cégep. Peu de Cégeps peuvent fournir un transport aux étudiants. L'adolescent est donc contraint à utiliser le transport en commun, mais est-ce un transport le plus approprié? Mettons-nous en situation, l'élève se rend à l'arrêt d'autobus comme à tous les matins, mais malheureusement l'autobus avait donc devancé son horaire et venait juste de partir, alors que fait-il? Il doit attendre le prochain passage et risque d'arriver en retard à ses cours. Cela est très désagréable et c'est pour cela que l'interdiction d'avoir son permis à 16 ans ne devrait pas être mis en vigueur.

Alors, juste pour cet argument le gouvernement devrait laisser tomber ce projet de loi qui ne convient vraiment pas à notre société. (Élève 2, paragraphes 3 et 4.)

Le début du troisième paragraphe annonce «quelques arguments», alors qu'il n'y a rien d'autre qu'un exemple sur lequel insiste de surcroît la conclusion. Cet exemple constitue le corps du texte, alors que des contre-arguments (qui ne sont jamais réfutés) ne sont que mentionnés en une seule ligne au second paragraphe.

### **L'interprétation des résultats**

Les items à réponse choisie mis à l'essai constituent une procédure indirecte de prise d'information, dont les résultats se rapportent à un aspect de la compétence langagière. Toutefois, le modèle de la compétence à écrire présenté dans la recherche comprend aussi, outre la compétence langagière, la compétence stratégique, qui permet à l'individu d'actualiser sa compétence langagière. Ce qui est observé dans les textes des élèves doit donc être interprété comme le fruit de l'interaction entre ces deux sous-composantes du modèle.

Ainsi, pour que les macrostratégies de cohésion sémantique soient observables dans le texte d'un élève, non seulement doit-il posséder l'habileté requise, c'est-à-dire cet aspect de la compétence langagière, mais aussi doit-il savoir s'en servir en situation de production de texte. Cette seconde opération relève de la compétence stratégique et s'effectue à travers une interaction entre le rédacteur et le lecteur éventuel et à travers les processus cognitifs d'écriture.

Dans les trois premières catégories, tous les textes présentent des tentatives d'actualisation des macrostratégies de cohésion sémantique. Ce

qui les distingue est que certains réussissent à le faire de façon efficace, et d'autres non. Ce qu'on observe dans les textes des élèves tend à indiquer que le score obtenu aux items est un indice de la maîtrise des macrostratégies de cohésion sémantique (voir tableau 2). Pour la première catégorie, tous les scores sauf un sont parmi les scores les plus élevés du groupe, et ainsi de suite jusqu'à la troisième catégorie, où les scores sont tous parmi les plus faibles du groupe. Le score des élèves dont les textes sont classés en (a) peut donc être interprété comme signifiant une maîtrise de la plupart des macrostratégies de cohésion sémantique, en (b) une maîtrise de certaines de ces stratégies et en (c) un manque de maîtrise des stratégies en question.

La seule exception à cette interprétation est la présence en (a) du texte d'un élève ayant obtenu, aux items à réponse choisie, le second plus faible score du groupe (25). Si on retourne à l'analyse de son texte, on se rend toutefois compte que les stratégies de cohésion qu'il a utilisées sont toutes très simples, quoique correctes, et font uniquement appel à la juxtaposition. Comme les items à réponse choisie faisaient souvent appel à des stratégies plus complexes d'organisation logique ou de discrimination de concepts similaires, on peut croire que ce candidat ne saurait employer ces stratégies, comme l'ont fait les autres candidats de cette catégorie. Cependant, il semble avoir développé suffisamment de compétence stratégique pour arriver à mettre en oeuvre des macrostratégies de cohésion sémantique minimales.

Que se passe-t-il en (d)? On peut arriver à y expliquer la présence d'élèves ayant obtenu des scores variés, en se servant à nouveau du modèle de la compétence à écrire. En effet, aucune tentative observable d'utilisation des macrostratégies de cohésion sémantique n'a pu être relevée dans leur texte. Il semblerait donc que, sans égard à leur maîtrise du construit mesuré (indiquée par le score aux items), ces élèves éprouvent tous certaines difficultés quant à la compétence stratégique. Ils n'ont pas tenté d'assurer la cohésion sémantique des structures d'ensemble de leur texte, même si les items semblent indiquer que certains possèdent l'habileté nécessaire. Éprouvent-ils des difficultés à planifier ou à réviser leur production, par exemple?

### **La validité des interprétations**

Les textes produits par les élèves ont permis d'interpréter les scores qu'ils ont obtenus aux items à réponse choisie, en les considérant les uns autant que les autres comme des indicateurs d'un construit, celui d'un aspect de la compétence à écrire tel que cette dernière est définie dans le

cadre conceptuel de la recherche. Les différentes relations au construit qui ont pu être mises en évidence à partir des données contribuent à la validité nomologique des interprétations. Cela ne signifie pas simplement qu'une représentation théorique rend directement valide les scores obtenus. Cela veut surtout dire que les interprétations des scores sont rendues plus crédibles dans la mesure où ceux-ci reflètent certaines implications théoriques du construit. Inversement, ce dernier devient aussi plus crédible, dans la mesure où les prédictions qu'il permet sont confirmées par des données empiriques (Messick, 1989).

Le premier construit auquel les données ont été reliées définit l'objet spécifique de la mesure effectuée par les prototypes d'items: les macro-stratégies de cohésion sémantique. Ce construit, perçu dans la présente recherche comme un aspect de la compétence langagière, est tiré des travaux de van Dijk et Kintsch (1983) sur les stratégies de compréhension du discours qui soutiennent que les mêmes stratégies sont requises lors de la production écrite.

Le second construit sur lequel est fondée la validité nomologique intègre le premier. En effet, ce second construit est le cadre de référence pour l'ensemble de la compétence à écrire. Le modèle adapté de Bachman (1990) semble pouvoir être utilisé de façon intéressante dans la conception d'instruments de mesure et l'interprétation des résultats obtenus. Son intérêt réside surtout dans la distinction qu'il propose entre la compétence langagière et la compétence stratégique, qui conduit à mieux comprendre la différence existant entre une procédure de mesure directe et une procédure de mesure indirecte ou, exprimé autrement, entre une mesure de la performance et une mesure de la compétence.

La confrontation dialectique entre les données empiriques et les théories, que favorise la validation nomologique, peut aussi conduire à la formulation d'énoncés provisoires susceptibles de contribuer à certains développements théoriques. Ainsi, les textes d'élèves analysés suggèrent que la compétence stratégique traite séparément la cohésion sémantique des paragraphes et celle du texte. On observe en effet, dans certains textes, une maîtrise assez bonne des stratégies qui concernent les paragraphes et une maîtrise plus faible de celles se rapportant au texte entier, ou inversement.

Les autres types d'évidences supportant la validité ne peuvent guère être allégués dans le cadre de la présente recherche. Pour établir la

représentativité de contenu, on a disposé du jugement d'experts en linguistique à propos de la congruence entre les tâches de mesure (les items mis à l'essai) et les macrostratégies de cohésion sémantique. Pour établir la fidélité structurelle, une analyse de la consistance interne du test que constituent les 39 items pourrait être effectuée (à l'aide de coefficients de corrélation item-test, par exemple). Cette consistance interne devrait être suffisamment grande, car le test mesure un construit relativement homogène. Dans le cas de construits plus englobants, il est normal d'observer une homogénéité réduite des réponses aux items. Idéalement, ce genre d'étude devrait être conduite auprès d'un plus grand groupe pour répondre aux exigences de l'analyse quantitative. Pour la validité en référence à un critère, il serait difficile de trouver un critère existant qui puisse être considéré pertinent. Les mesures directes, selon le modèle de la compétence à écrire préconisé, constitueraient de faibles critères d'un aspect de la compétence langagière à cause, on l'a souligné, de l'interaction entre les compétences langagière et stratégique. On pourrait élaborer d'autres items et les mettre à l'essai, puis calculer la corrélation entre les scores ainsi obtenus et ceux de la présente recherche, chaque mesure servant en quelque sorte de critère à l'autre. Cette approche relève toutefois davantage de la portée nomothétique, car elle procède par l'accumulation de résultats convergents entre recherches.

La validation des interprétations peut aussi comprendre une analyse des conséquences de l'utilisation de ces dernières (la validité écologique des interprétations). Dans la présente recherche, les interprétations des scores ne sont pas utilisées pour quelque intervention pédagogique que ce soit auprès des élèves, mais uniquement pour analyser leur validité de construit. Leur utilisation pourrait toutefois sembler justifiée à des fins diagnostiques et de remédiation. À cet effet, un faible score indiquerait une lacune dans la maîtrise des MCS, et c'est surtout à développer cet élément de compétence que devrait s'employer le candidat. Un score élevé, conjugué à une absence de stratégie observable dans le texte, signifierait que la compétence stratégique fait défaut, et c'est surtout à développer cette compétence (la maîtrise des processus cognitifs d'écriture, par exemple) que le candidat aurait à travailler. Les différents sous-scores pourraient guider davantage ce dernier dans la découverte des lacunes, puisque les séries d'items abordaient la cohésion sémantique tantôt sous l'angle de la planification, tantôt sous l'angle de la révision.

La présente recherche suggère qu'il est possible de mesurer avec une certaine validité un aspect de la compétence à écrire autre que la maîtrise du code (orthographe, vocabulaire, syntaxe) en utilisant des items à réponse

choisie. Des recherches ultérieures pourraient tenter d'analyser la validité des interprétations faites à partir de scores obtenus à propos d'autres aspects de la compétence à écrire, soit sur le plan de la compétence langagière ou de la compétence stratégique, à l'aide de tels items.

## Conclusion

Les résultats rapportés nous semblent stimulants et soulèvent avec plus de pertinence encore le questionnement qui a été à l'origine de la présente recherche. On a pu analyser une qualité importante de la validité des interprétations faites à partir des items à réponse choisie mis à l'essai: leur validité nomologique. Que doit-on en conclure? Inévitablement, à cause de la portée limitée de ces résultats, que d'autres travaux sur le sujet semblent justifiés et seraient nécessaires afin d'apporter une réponse plus complète. Cette conclusion peut paraître circonspecte, mais est en accord avec la nature exploratoire de la démarche. Afin d'orienter d'autres investigations sur le sujet, on peut formuler deux remarques à propos des choix méthodologiques qui ont été privilégiés.

Premièrement, la présente recherche a permis d'illustrer la nécessité, afin de comprendre les phénomènes d'écriture et la mesure de la compétence à écrire, de recourir à un cadre conceptuel qui distingue les habiletés langagières d'écriture des habiletés stratégiques nécessaires à leur actualisation lors d'une production écrite. On peut être relativement habile à assurer la cohésion sémantique d'un paragraphe et, comme on l'a vu, rédiger néanmoins des paragraphes dont la cohésion sémantique est insuffisante.

Deuxièmement, l'utilité d'une approche qualitative pour une recherche sur la validité des mesures de la compétence à écrire a été mise en évidence. Les observations tirées de l'analyse de cas ont permis de décrire certains des phénomènes en jeu et de rendre compte de leur diversité.

La contribution la plus importante de cette recherche, au-delà de ses résultats ou des conclusions qu'on en tire, est peut-être l'esprit qui l'anime: l'insistance sur la recherche de la validité des interprétations qui sont faites à partir de mesures en éducation. L'importance, autant pour les individus que pour les populations, des décisions qui se fondent sur ces interprétations exige que cette préoccupation se traduise en efforts concrets de validation de la part des concepteurs d'instruments de mesure et de leur utilisateurs.

## RÉFÉRENCES

- Ackerman, T.A. & Smith, P.L. (1988). A comparison of the information provided by essay, multiple-choice and free-response writing tests. Applied Psychological Measurement, 12(2), 117-128.
- American Psychological Association, American Educational Research Association et National Council on Measurement in Education (APA) (1966, 1974, 1985). Standards for educational and psychological testing. Washington : American Psychological Association.
- Angoff, W. H. (1988). Validity: a evolving concept. In H. Wainer & H. Braun (Éds), Test validity (pp. 19-32). Hillsdale: Lawrence Erlbaum Associates.
- Asselin, C. & Mc Laughlin, A. (1992). La rédaction comme moyen d'évaluer la compétence en français écrit. Pédagogie collégiale, 5( 3), 11-15.
- Bachman, L.F. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.
- Benton, S.L. & Kiewra, K.A. (1986). Measuring the organizational aspects of writing ability. Journal of Educational Measurement, 23 (4), 377-386.
- Bereiter C. & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, N.J. : Lawrence Erlbaum Associates.
- Bibeau, G., Lessard, C., Paret, M.-C. & Thérien, M. (1987). L'enseignement du français langue maternelle: perceptions et attentes. Québec: Gouvernement du Québec, Conseil de la langue française, collection Dossiers, n° 27.
- Blais, J.-G. & Auger, R. (1994). Une stratégie d'estimation de la fidélité inter-correcteurs pour une épreuve de rédaction dans une enquête internationale. Communication présentée à l'occasion de la 17<sup>e</sup> session d'étude de l'Association pour le développement de la mesure et de l'évaluation en éducation, Québec, octobre.
- Braun, H.I. (1988). Understanding scoring reliability: experiment in calibrating essay readers. Journal of Educational Statistics, 13(1), 1-18.
- Breland, H.M., Camp, R., Jones, R.J., Morris, M.M. & Rock, D.A. (1987). Assessing writing skill. New York: College Entrance Examination Board.
- Canale, M. (1983). On some dimensions of language proficiency. In J.W. Oller (Ed.), Issues in language testing research (pp. 333-342). Rowley: Newbury House.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1(1),1-47.
- Charney, D. (1984). The validity of using holistic scoring to evaluate writing: a critical overview. Research in the Teaching of English, 18(1), 65-81.
- Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge: MIT Press.

- Conlan, G. (1986). Objectives measures of writing ability. In K.L. Greenberg, H.S. Wiener & R.A. Donovan (Éds), Writing assessment: issues and strategies. New York: Longman.
- Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Éd.), Educational measurement (2<sup>e</sup> éd.) (pp. 443-507). New York: American Council on Education.
- Cronbach, L.J. (1988). Five perspectives on validity argument. In H. Wainer & H.I. Braun (Éds), Test validity, (pp. 3-17). Hillsdale: Lawrence Erlbaum Associates.
- Fagnant, G. & Rousseau, C. (1990). Élaboration et expérimentation d'une stratégie d'intervention en français écrit au collégial. Rapport de recherche, Association des collèges du Québec.
- Ferland, L.-G. & Sanchez, E. (1990). Tests de français à l'entrée des collèges et des universités. Québec: Gouvernement du Québec, Direction générale de l'enseignement collégial.
- Flower, L.S. & Hayes, J.R. (1981). Plans that guide the composing process. In C.H. Frederiksen & J.F. Dominic (Éds), The nature, development, and teaching of written communication (vol.2, pp. 39-58). Hillsdale: Lawrence Erlbaum Associates.
- Frederiksen, C.H. & J.F. Dominic (1981). Perspectives on the activity of writing. In C.H. Frederiksen & J.F. Dominic (Éds), The nature, development, and teaching of written communication (vol. 2, pp. 1-20). Hillsdale: Lawrence Erlbaum Associates.
- Gouvernement du Québec (1986). Les résultats de l'épreuve de français écrit de cinquième secondaire administrée au mois de mai 1986. Québec: Ministère de l'Éducation.
- Haladyna, T.M. (1992). The effectiveness of several multiple-choice formats. Applied Measurement in Education, 5(1), 73-88.
- Hayes, J.R. & Flower, L.S. (1980). Identifying the organization of writing processes. In E.W. Gregg & E.R. Steinberg (Éds), Cognitive processes in writing (pp. 3-29). Hillsdale: Lawrence Erlbaum Associates.
- Hymes, D. (1974). Foundations in sociolinguistics. Philadelphie: University of Philadelphia Press.
- Lebrun, M. (1987). Vers un modèle intégré des critères de compréhension en lecture au collégial. Thèse de doctorat non publiée, Université Laval, Sainte-Foy.
- Lecavalier, J. (1991). Mesure et démesure de la compétence. Factuel, 4(1), 7-9.
- Lefebvre, J. & Turcotte, A.G. (1989). Éléments pour une définition de la compétence langagière propre au collégial : rapport d'étude. Québec: Ministère de l'Enseignement supérieur et de la Science, Sous-comité pédagogique, Direction générale de l'enseignement collégial.



- Messick, S. (1989). Validity. In R.L. Linn (Éd.), Educational measurement (3<sup>e</sup> éd., pp. 13-103). New York: American Council on Education.
- Miller, M.D. & Crocker, L. (1990). Validation methods for direct writing assessment. Applied Measurement in Education, 3(3), 285-296.
- Nold, E.W. (1981). Revising. In C.H. Frederiksen & J.F. Dominic (Éds) The nature, development, and teaching of written communication (vol. 2, pp. 67-79). Hillsdale: Lawrence Erlbaum Associates.
- Oller, J.W. (1978). How important is language proficiency to IQ and other educational tests? In J.W. Oller & K. Perkins (Éds), Language in education: testing the tests (pp. 1-16). Rowley: Newbury House Publishers.
- Oller, J.W. (1983). Evidence for a general language proficiency factor: an expectancy grammar. In J.W. Oller (Éd.), Issues in language testing research (pp. 3-10). Rowley: Newbury House Publishers.
- Stiggins, R.J. (1982). A comparison of direct and indirect writing assessment models. Research in the teaching of English, 16(2), 101-114.
- Van Dijk, T.A. & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.