

L'influence de la compréhension en lecture sur les résultats aux tests utilisés en éducation

Serge P. Séguin

Volume 7, Number 1, Winter 1981

URI: <https://id.erudit.org/iderudit/900315ar>

DOI: <https://doi.org/10.7202/900315ar>

[See table of contents](#)

Publisher(s)

Revue des sciences de l'éducation

ISSN

0318-479X (print)

1705-0065 (digital)

[Explore this journal](#)

Cite this article

Séguin, S. P. (1981). L'influence de la compréhension en lecture sur les résultats aux tests utilisés en éducation. *Revue des sciences de l'éducation*, 7(1), 37–46. <https://doi.org/10.7202/900315ar>

Article abstract

Several studies demonstrate that reading comprehension constitutes an important variable measured by tests used in education. This presents the problem of interpretation of results, whether research results or scholastic achievement scores, obtained from tests designed to measure cognitive or affective variables. Solutions to the problem of accurate measurement are proposed and discussed, in the context of evaluative decisions the quality of which depends on the validity of the data collected.

L'influence de la compréhension en lecture sur les résultats aux tests utilisés en éducation¹

Serge P. Séguin*

Résumé — Plusieurs études démontrent que la compréhension en lecture constitue une variable importante mesurée par les tests utilisés en éducation. Cette situation pose un problème d'interprétation des résultats, qu'il s'agisse de résultats de recherche ou de résultats scolaires, obtenus à des épreuves destinées à mesurer des variables cognitives ou affectives. Des voies de solution à ce problème de pertinence de la mesure sont proposées et discutées, dans le cadre de décisions évaluatives dont la qualité dépend de la validité de l'information recueillie à l'aide d'instruments de mesure.

Abstract — Several studies demonstrate that reading comprehension constitutes an important variable measured by tests used in education. This presents the problem of interpretation of results, whether research results or scholastic achievement scores, obtained from tests designed to measure cognitive or affective variables. Solutions to the problem of accurate measurement are proposed and discussed, in the context of evaluative decisions the quality of which depends on the validity of the data collected.

Resumen — Varios estudios demuestran que la comprensión de la lectura constituye una variable importante medida en las pruebas utilizadas en educación. Esta situación plantea un problema en la interpretación de los resultados, bien sea en los resultados de la investigación o en los resultados escolares, obtenidos en pruebas destinadas a medir las variables cognitivas o afectivas. Algunas vías de solución a este problema relacionado con lo apropiado de la medida son propuestos y discutidos en el marco de decisiones evaluativas donde la calidad depende de la validez de la información recogida mediante instrumentos de medida.

Zusammenfassung — Mehrere Untersuchungen beweisen, dass das Verständnis beim Lesen eine wichtige Variable darstellt, die man durch in der Pädagogik verwendete Tests misst. Diese Situation stellt ein Problem bezüglich der Interpretation der Resultate, sei es bei Forschungs-, sei es bei schulischen Ergebnissen, die man durch Prüfungen erhält, welche auf intellektuelle oder affektive Variable ausgerichtet sind. Im Rahmen von Bewertungsentscheidungen, deren Qualität von der Gültigkeit der mithilfe von Messwerkzeugen gesammelten Information abhängt, werden Lösungswege für dieses Problem des Zutreffens der Bemessung vorgeschlagen und untersucht.

Mesure et décisions évaluatives en éducation

Tout instrument de mesure utilisé en éducation ne peut être valable que s'il permet de supporter des décisions évaluatives qui soient le plus possible conformes à la réalité concernée (Séguin, 1979). Ces décisions peuvent être plus ou moins complexes et, dans une situation donnée, répondre à des préoccupations d'un seul ou de plusieurs ordres. Sans

* Séguin, Serge P. : professeur, Université du Québec à Montréal.

prétendre en épuiser la nomenclature, il s'avère utile d'identifier trois ordres majeurs de décisions évaluatives en éducation, selon que ces décisions répondent à des préoccupations pédagogiques, administratives ou investigatrices.

Les décisions évaluatives peuvent être de nature *pédagogique ou didactique* : les mesures obtenues constituent alors des éléments d'information permettant d'orienter l'action pédagogique et l'intervention didactique du maître. Les décisions évaluatives peuvent aussi être d'ordre *administratif* : les mesures obtenues constituent alors des éléments d'information permettant de décrire certaines caractéristiques des sujets mesurés en vue d'orienter l'action des administrateurs. Les décisions évaluatives peuvent enfin concerner des activités de *recherche* : les mesures obtenues constituent en ce cas des éléments d'information, s'ajoutant à d'autres informations déjà connues, soit pour élargir le champ des connaissances, soit pour approfondir la connaissance d'un phénomène particulier soumis à une étude systématique.

Dans chacun de ces trois ordres de préoccupations, la qualité de la décision évaluative repose notamment sur la compétence de l'évaluateur (enseignant, administrateur, chercheur) et sur la validité de l'information recueillie.

Validité de l'information

Qu'est-ce qu'une information valide ? Une réponse peut être trouvée en parcourant l'histoire du développement des connaissances en testing dans les domaines de l'éducation, de la psychologie, et dans toutes les autres sciences du comportement humain. Toute cette évolution pourrait être focalisée autour d'une préoccupation majeure : la recherche constante des meilleurs moyens d'obtenir les informations les plus valides en regard des décisions à prendre. Dans la littérature, on ne retrouve que très rarement l'expression « information valide ». Certains auteurs semblent rattacher le concept de validité exclusivement aux scores d'un test, c'est-à-dire aux seuls résultats d'un instrument de mesure (*Standards...*, 1974 ; Guion, 1977). Questionner la validité, c'est alors se demander ce qui peut être adéquatement inféré à partir des scores obtenus ; la validité concerne soit ce que le score observé représente en fait, soit l'inférence d'un score probable à une autre mesure. Dans un sens plus large, Cronbach (1971) parle de la validité d'un test comme recouvrant à la fois la validité des scores (de leur interprétation) et de l'utilisation qu'on en fait. C'est cette dualité que recouvre le concept de validité de l'information.

Valider une information, c'est en investiguer systématiquement les processus de cueillette et d'utilisation, dans le but de démontrer la valeur de cette information à l'appui d'une décision évaluative. Cette investigation concerne principalement :

- la *représentativité* de ce qui est mesuré,
- la *relation* de ce qui est mesuré avec d'autres mesures,
- la *précision* avec laquelle la mesure est obtenue, et
- la *pertinence* de la mesure en regard des décisions à prendre.

Investiguer la représentativité de ce qui est mesuré constitue un processus généralement associé à la validation de contenu du test (Ebel, 1972 ; Guion, 1977). Apprécier la relation de ce qui est mesuré avec d'autres mesures relève de la validité corrélacionnelle ou de construit du test, alors que la précision des scores concerne leur fidélité (*Standards...*, 1974). Dans la suite de nos propos, nous traiterons plus particulièrement de la *pertinence* de la mesure (de l'information recueillie).

Pertinence de la mesure

Le terme *pertinence* désigne ici la qualité de mesurer réellement et seulement ce qu'on désire mesurer. Morissette (1979) utilise ce terme dans un sens différent : il parle de « pertinence empirique » (p. 250) pour désigner un processus de validation corrélacionnelle, et de « pertinence logique » (p. 252) pour désigner la congruence entre les questions d'un test de rendement et les objectifs mesurés. Bien que le concept de pertinence retenu ici peut se rapprocher du concept de « pertinence logique » de Morissette dans le cas d'examens de rendement, il est ici utilisé comme l'équivalent du concept de « relevance » décrit par Ebel (1971, p. 362-364), à savoir : l'ensemble des spécifications et des prescriptions (types d'items, formats de réponses, directives, conditions, etc.) qui font que la réussite à chaque item du test reflète une compétence réelle du sujet sur l'attribut mesuré (un trait, une connaissance, une habileté, un domaine, etc.) et non sur autre chose.

La mesure doit donc être pertinente si l'on veut que l'information recueillie soutienne des décisions évaluatives efficaces et conformes à la réalité. Ainsi, si une épreuve est destinée à mesurer le rendement dans une matière scolaire, il faut que cette épreuve permette de le bien mesurer ; si une autre épreuve est bâtie pour mesurer l'attitude des enseignants vis-à-vis l'utilisation d'objectifs de comportement, il faut que les scores obtenus reflètent bien l'attitude des enseignants, et non une autre caractéristique. La pertinence d'un instrument de mesure requiert non seulement que la variable mesurée soit définie opérationnellement, mais aussi que les composantes de la mesure soient précisées. Ces deux conditions sont nécessaires, bien que non suffisantes, pour pouvoir bien interpréter les résultats obtenus.

Par exemple, que mesure précisément un test d'orthographe administré sous forme de dictée verbale ? En terme de contenu, un tel test mesure l'habileté à écrire correctement des mots dictés. Mais encore ? Cronbach (1971, p. 453) précise qu'un tel test est susceptible de mesurer à la fois des habiletés de discrimination auditive, d'orthographe, de vocabulaire et d'écriture. Un tel test manquerait donc de pertinence si l'on voulait mesurer uniquement des connaissances en orthographe.

Il existe une foule d'autres exemples qui pourraient être cités. Notre intention est de faire remarquer que certaines habiletés en lecture sont souvent des préalables dans la plupart des instruments de mesure utilisés en éducation. Cela pose naturellement un problème majeur concernant la pertinence des mesures obtenues à l'aide de ces instruments. Ce problème concerne directement l'interprétation de ce qui est

effectivement mesuré et est susceptible d'affecter soit la validité des décisions pédagogique, soit la validité des conclusions de nombreuses recherches en éducation.

Plusieurs études ont investigué l'influence d'habiletés linguistiques sur les résultats à des tests de toutes sortes utilisés en éducation. Pour des raisons qui deviendront évidentes plus loin, il est utile de distinguer ici les études qui ont exploré la présence dans les tests scolaires d'un facteur général relié à la maîtrise linguistique, et les études qui ont exploré d'une façon plus particulière la détermination d'habiletés en lecture sur les résultats à des tests de rendement scolaire. Ces deux catégories d'études sont résumées dans les deux sections suivantes.

La maîtrise de la langue : un facteur général

Depuis plusieurs années, nombre de chercheurs ont investigué l'influence de la maîtrise générale de la langue (« language proficiency ») sur les résultats à divers tests (d'intelligence, de rendement, de personnalité) utilisés en éducation. Résumant les études menées en ce sens, et proposant d'autres points d'exploration, Oller et Perkins (1978) démontrent que la maîtrise de la langue constitue la principale variable mesurée par de nombreux tests. La maîtrise de la langue apparaît comme un facteur général susceptible même d'être confondu avec un facteur général d'intelligence. Pour ces auteurs, comme pour leurs collaborateurs (Gunnarsson, 1978 ; Streiff, 1978 ; Stump, 1978), cela n'a rien d'étonnant puisque : 1) la langue constitue le véhicule de l'enseignement et le support aux directives des tests ; 2) la plupart des caractéristiques de l'intelligence sont identiques à celles utilisées par l'être humain pour apprendre et utiliser une langue ; 3) les processus mentaux requis pour traiter une information, linguistique ou autre, sont les mêmes ; et 4) la langue constitue un facteur-clé pour la scolarisation, de sorte que l'expérience d'apprentissage de la langue apparaît comme un processus naturel de sélection entre les élèves.

La maîtrise de la langue est définie dans ces études, notamment par Stump (1978), comme une habileté à opérer dans cette langue, habileté dont le noyau central consiste à intégrer des composantes linguistiques (v.g. : phonologie, syntaxe, etc.) et diverses compétences (lecture, écriture, écoute, parole).

Les conclusions de Oller et Perkins (1978) reposent sur des études qui comparent les résultats à divers tests couramment utilisés en éducation avec les résultats obtenus à des tests destinés à mesurer la maîtrise de la langue (« language proficiency tests »). Leurs conclusions suggèrent que, dans l'interprétation des scores observés à divers tests scolaires, l'on tienne compte de la part déterminée par un facteur général de maîtrise de la langue.

Mais existe-t-il aussi, dans divers tests, un facteur plus spécifique relatif au savoir-lire des répondants ? Existe-t-il un facteur associé à des habiletés en lecture et dont l'effet sur les résultats aux tests pourrait être circonscrit, contrôlé, ou même annulé ?

Présence d'habiletés en lecture

On sait depuis longtemps que les épreuves de rendement scolaire reposent plus ou moins directement sur des habiletés en lecture (« reading abilities »). La présente section résume quelques études portant sur cette relation. Nous ne connaissons aucune étude qui fasse mention de la présence possible d'habiletés en lecture dans des instruments destinés à mesurer des attitudes, des intérêts, ou d'autres attributs ou traits affectifs. Toutefois, les études résumées ici laissent croire que des habiletés en lecture sont aussi présentes lorsqu'il s'agit de mesurer des variables affectives. L'expérience démontre d'ailleurs que le répondant à un item se demande souvent ce que le rédacteur a bien voulu signifier exactement.

Nous avons retenu ici trois études récentes qui démontrent que les tests de rendement (en sciences de la nature, en mathématiques et en sciences de la santé) seraient grandement tributaires des habiletés en lecture des répondants.

Séguin et Bélair-Pépin (1978) ont étudié diverses variables susceptibles d'influencer l'efficacité d'un apprentissage des sciences au secondaire par objectifs opératoires, selon un curriculum mis au point et expérimenté par le GREC², avec la collaboration de nombreux enseignants de plusieurs commissions scolaires du Québec. Onze programmes en sciences en vigueur au Québec ont ainsi été développés, et cinq d'entre eux ont été expérimentés, dans trois écoles, totalisant 13 groupes-cours regroupés en 7 programmes-écoles et définissant autant de groupes expérimentaux. L'étude portait donc sur 470 sujets répartis entre 7 groupes expérimentaux comprenant chacun de 21 à 122 sujets. Sept variables indépendantes ont été étudiées pour expliquer le rendement de ces élèves en sciences. Les coefficients de corrélation linéaire entre le rendement en sciences et chacune des 7 variables indépendantes ont été calculés dans chaque groupe. On retrouve notamment des coefficients de corrélation variant de 0,29 à 0,56 entre le rendement en sciences et le résultat à un test portant sur la compréhension en lecture d'un texte à contenu scientifique.

En utilisant un modèle linéaire de régression multiple, les auteurs concluent que le succès en sciences des sujets expérimentaux était généralement lié à la maîtrise des prérequis mathématiques, puis à la compréhension en lecture, et en troisième lieu au quotient intellectuel.

Ces résultats confirmaient alors une conclusion de Desautels et Legendre (1971) à l'effet que la compréhension en lecture ferait partie des deux ou trois facteurs les plus prédictifs de la réussite scolaire en physique au niveau secondaire. Cette dernière étude comprenait un échantillon représentatif de 1015 adolescents. Elle portait sur l'impact de 15 variables susceptibles d'être significatives des apprentissages en sciences exactes. La variable « compréhension de lecture » s'est révélée tout aussi importante, quelle que soit la méthode d'apprentissage utilisée par les sujets expérimentaux.

Dans une étude plus récente, Benson et Crocker (1979) démontrent que l'habileté en lecture a un effet significatif sur le rendement à un test en sciences de la santé au niveau secondaire. Les 86 sujets expérimentaux ont été divisés en deux groupes d'habiletés en

lecture, selon le score à un test général de lecture. Selon les normes nationales américaines pour ce test, les sujets expérimentaux ont obtenu des scores en lecture qui les situaient en dessous du 82e percentile : les sujets ont donc été divisés en « moyens » et en « faibles » selon leur habileté en lecture. Les « moyens » en lecture (N = 40) ont réussi deux fois mieux, aux 72 items contenus dans les tests de rendement, que les « faibles » en lecture (N = 46), peu importe qu'il s'agisse d'items du type vrai-ou-faux, choix multiples ou association.

Dans une autre étude, Landry (1979) rapporte un coefficient linéaire de 0,54 entre les résultats de 45 élèves du niveau secondaire I à un test de rendement en mathématiques composé de 20 items du type choix multiples et leurs scores à un test de compréhension en lecture, construit selon une forme de closure et dans lequel le répondant devait compléter 45 lacunes (trouver les 45 mots manquants) d'un texte. Ce test de compréhension en lecture a été validé en 1975-76 auprès d'un échantillon représentatif des élèves québécois du niveau secondaire I. Quant aux 20 items du test de rendement en mathématiques, 18 présentaient des choix numériques de réponses, et un autre de figures géométriques, alors qu'un seul présentait cinq phrases d'une ligne chacune parmi lesquelles il fallait choisir l'énoncé qui était faux. Quant aux énoncés des questions, la moitié n'occupait qu'une ligne et un seul occupait plus de deux lignes.

Ces trois études paraissent confirmer que la compréhension en lecture influence généralement d'une façon significative la performance à un test scolaire. Les habiletés en lecture, concernées dans ces dernières études, apparaissent relativement spécifiques : on ne peut prétendre qu'elles ne constituent que des mesures d'une certaine intelligence générale. En effet, d'une part, dans l'étude de Séguin et Bélair-Pépin (1976), les coefficients de corrélation linéaire entre le rendement en sciences et la compréhension en lecture étaient sensiblement supérieurs aux coefficients entre ce rendement et l'habileté mentale générale (environ 0,42 contre 0,29 en moyenne). D'autre part, comme il en sera fait mention plus loin (voir Section 7), Landry (1979) a démontré que le coefficient de corrélation linéaire entre le rendement et la compréhension en lecture subit une baisse abrupte lorsque les questions sont lues à haute voix aux mêmes élèves.

En outre, on ne peut prétendre que les épreuves utilisées dans ces trois études pour mesurer le savoir-lire ne mesuraient que la maîtrise de la langue telle que définie en section 4. En effet, d'une part, Benson et Crocker (1978) ont utilisé un test d'habiletés en lecture (« reading ability ») destiné à mesurer un domaine relativement restreint de la maîtrise générale de la langue (« language proficiency »). D'autre part, dans les deux études québécoises précédentes, des tests lacunaires ont été utilisés à cette fin. Cette forme de test a été retenue parce que sa validité est généralement reconnue pour mesurer la « compréhension » en lecture, et parce que Oller (voir Streiff, 1978 ; aussi De Landsheere, 1975) a déjà démontré que les coefficients varient généralement d'environ 0,70 à 0,95 entre les résultats à des tests lacunaires et les résultats à d'autres tests mesurant l'habileté en lecture. Il est certain que ces tests mesurent aussi un facteur général, identifiable à l'intelligence ou à la maîtrise de la langue. Mais l'important est ici de noter que ces tests mesurent aussi des habiletés plus spécifiques relatives au savoir-lire. Il apparaît donc

possible d'isoler le savoir-lire dans la détermination des résultats à divers tests de rendement scolaire.

Le savoir-lire semble relié plus spécifiquement à la compréhension des textes écrits, faisant appel au décodage des mots et au sens d'une phrase ou d'un texte. Dans la suite des présents propos, la pertinence de ce facteur compréhension en lecture retient l'attention.

Pertinence ou non des habiletés de compréhension en lecture

La validité d'un test doit être déterminée en considérant comme pertinente ou non la présence d'habiletés linguistiques dans le domaine mesuré par ce test. Si l'on juge ces habiletés pertinentes et qu'elles sont présentes dans la mesure, il faut le préciser et en tenir compte dans l'interprétation et l'utilisation des résultats. Si ces habiletés sont jugées impertinentes et qu'elles sont présentes dans la mesure, il faut alors les isoler, et les extraire ; sans quoi l'interprétation et l'utilisation des résultats risquent d'invalider les décisions prises.

Dans cette perspective, la présence d'un facteur général, qu'il s'agisse d'intelligence ou de maîtrise de la langue telle que définie en section 4, ne saurait être reniée, ni même jugée impertinente. Tout au plus conviendra-t-il d'en tenir compte dans l'interprétation et dans l'utilisation des résultats.

Quant à la pertinence ou non des habiletés de compréhension en lecture, elle doit être jugée en regard des buts du test, c'est-à-dire en regard des décisions qui seront prises sur la base de l'information recueillie (Guion, 1977) : si ces habiletés sont impertinentes au domaine concerné, l'instrument de mesure et l'information recueillie seraient d'autant moins valides que ces habiletés seraient variables dans le groupe des sujets mesurés.

Les habiletés de compréhension en lecture se révèlent être des parasites à l'intérieur des instruments destinés à mesurer la performance des élèves dans diverses matières scolaires. Un problème majeur doit alors être résolu dans la construction de tels instruments si l'on veut minimiser l'effet indésirable causé par la présence, en degrés variables chez les répondants, d'habiletés spécifiques en compréhension de lecture.

Une voie de solution ?

Pour réduire au minimum l'effet parasitaire d'habiletés en lecture sur la performance aux instruments de mesure utilisés en éducation, Gardner (1978) propose de simplifier la syntaxe et le vocabulaire employés. Le vocabulaire spécifique au domaine concerné doit être utilisé pour bien mesurer ce domaine, mais la performance ne doit pas dépendre d'habiletés générales en lecture.

Cette approche de simplification est-elle satisfaisante ? Sans doute qu'elle devrait permettre d'augmenter la pertinence d'un test destiné à mesurer un domaine spécifique. Néanmoins, dans l'étude de Benson et Crocker (1979), l'effet d'habiletés en lecture s'est avéré très significatif malgré l'effort des auteurs pour simplifier la syntaxe et le vocabulaire de leurs tests de rendement en sciences de la santé. De même, dans l'étude de Landry

(1979), la syntaxe et le vocabulaire employés dans le test de rendement en mathématiques ont été simplifiés à l'extrême ; pourtant, cela n'a pas empêché que près de 30% de la variance observée dans la performance à ce test soit explicable par la variance des scores à un test de compréhension en lecture.

Pour contrer cet effet indésirable du facteur compréhension en lecture, plusieurs enseignants semblent avoir pris l'habitude de lire les questions à haute voix à leurs élèves, en leur laissant une courte période pour répondre à chaque question qui vient d'être lue. Landry (1979) a expérimenté ce procédé. Deux effets de ce procédé méritent d'être soulignés ici. D'une part, le coefficient de corrélation linéaire entre le rendement au test de mathématiques et la performance au test de compréhension en lecture est passé de 0,54 (voir section 5) à -0,32 lorsque les questions étaient lues à haute voix par l'enseignant : la proportion de variance commune est donc passée d'environ 30% à moins de 10%. Cela signifie que les différences observées entre les sujets au test de mathématiques étaient attribuables à 30% à leurs différences dans la compréhension des directives écrites. Lorsque les directives ont été lues à haute voix, plusieurs différences dans la compréhension de ces directives ont été aplanies, de sorte que les différences restantes n'expliquaient plus que 10% des différences observées en mathématiques. D'autre part, la réussite au test lu à haute voix a été significativement supérieure, au seuil de 0,001, à la réussite au test administré d'une façon conventionnelle, la réussite moyenne du groupe passant de 79% à 87%.

Il faut noter cependant que, dans l'étude de Landry, les items qui ont été significativement mieux réussis au test lu à haute voix comprennent : 1) un item qui mesure la compréhension du concept de fonction, 2) deux items qui requièrent des connaissances terminologiques très précises, et 3) les six items mesurant l'habileté à effectuer des opérations de multiplication et de division. Il semble donc que les élèves les plus faibles en compréhension en lecture en général sont aussi ceux qui présentent le plus de difficultés en « lecture fonctionnelle », c'est-à-dire dans la compréhension des directives écrites concernant un domaine particulier du savoir, ici les mathématiques. Pour ces élèves, la signification du mot écrit paraît difficile mais un sens est associé plus facilement à ce mot entendu.

La présence dans les scores en mathématiques d'habiletés relatives à la compréhension de textes (ici des directives ou des questions) n'a évidemment pas été éliminée : c'est un préalable qui fait probablement partie du facteur général de maîtrise de la langue nécessaire à la réussite d'une majorité de tests utilisés en éducation. Mais, par l'écoute des directives, la compréhension s'accroît chez la plupart des répondants qui n'arrivaient pourtant pas à bien saisir les messages présentés uniquement par écrit. Ce qui importe, ce n'est pas que les scores en mathématiques soient plus élevés, ni que le test devienne plus facile ; le plus important est que les scores en mathématiques reflètent plus directement des performances relatives à ce domaine et non à un amalgame nébuleux de mathématiques et de compréhension de textes écrits.

Plusieurs autres études devront encore être poursuivies avant de pouvoir conclure d'une façon plus définitive. L'étude de Landry (1979) confirme toutefois que le double

support (directives écrites et lues, ou mieux : lues et entendues) favorise la compréhension des directives, tout comme ce double support améliore tout apprentissage. Cette étude paraît aussi soutenir la théorie des styles cognitifs à l'effet que certains apprenants se sentiraient plus à l'aise dans un contexte plus visuel qu'auditif ou plus auditif que visuel (Hill, 1974).

L'étude de Landry suggère un moyen de rendre presque négligeable l'influence que l'incompréhension des directives peut avoir sur la performance mesurée par divers tests utilisés en éducation, principalement par les tests de rendement. Qu'il s'agisse de tests standardisés ou non, il faut toujours s'assurer que le répondant comprenne bien les directives. Ainsi, on pourra apprécier, non pas sa compréhension des directives, mais sa compétence à résoudre les situations-problèmes qu'on lui présente.

Conclusion

Une suggestion d'ordre pédagogique découle des constatations précédentes. Il semble en effet que les enseignants de diverses disciplines seraient bien avisés d'insister sur la pratique de la lecture silencieuse, afin d'habituer les élèves à comprendre les consignes écrites particulières à différents domaines du savoir. Il serait utopique de s'attendre à ce que toutes les déficiences en lecture soient corrigées. L'enseignement doit se poursuivre et l'information pertinente aux décisions évaluatives d'ordre pédagogique, didactique, administratif, ou de recherche doit continuer d'être recueillie.

Afin de mieux adapter la présentation des instruments aux possibilités des répondants et obtenir ainsi des mesures plus pertinentes, une solution est de simplifier la syntaxe et le vocabulaire employés. Dans le domaine des tests de rendement scolaire, il est de plus suggéré que les questions soient lues à haute voix aux élèves. Toutefois, l'étude de Landry (1979) suggère qu'un tel procédé peut introduire une autre forme d'impertinence si la période de temps pour répondre était ainsi trop écourtée : il ne faut pas que le facteur « vitesse d'exécution » soit substitué au facteur « compréhension en lecture ».

En résumé, lorsqu'on mesure la performance d'élèves dans des domaines précis, il faut savoir cerner la part attribuable aux habiletés en lecture, et même réduire sinon annuler cet effet parasite s'il devient nuisible aux décisions à prendre. Il en est de même de toute épreuve destinée à mesurer des attitudes et exigeant du répondant qu'il indique son accord ou son évaluation d'un énoncé présenté : il faut alors s'assurer que tout répondant comprenne parfaitement bien l'énoncé.

NOTES

1. L'auteur désire remercier les participants à un séminaire tenu sur le sujet à la Faculté des sciences de l'éducation de l'Université de Montréal le jeudi 24 janvier 1980. En outre, l'auteur exprime ses remerciements au Dr Renald Legendre, professeur au Département des sciences de l'éducation de l'Université du Québec à Montréal, pour ses commentaires sur une version préliminaire de cet article.
2. GREC : Groupe de recherche en évaluation des curriculum, Département des sciences de l'éducation, Université du Québec à Montréal.

RÉFÉRENCES

- Benson, J. et Crocker, L., The effects of item format and reading ability on objective test performance : a question of validity, *Educational and Psychological Measurement*, 39, 1979, p. 381-387.
- Cronbach, L.J., Test validation, In R.L. Thorndike (Ed.) : *Educational Measurement* (2ème éd.), Washington, D.C. : American Council on Education, 1971, ch. 14.
- DeLandsheere, G, *Le test de closure*. Paris : Fernand Nathan, 1973.
- Désautels, J. et Legendre, R, *Étude expérimentale comparative des principaux procédés pédagogiques utilisés dans l'enseignement de la physique au Québec*, Institut de Recherche Pédagogique, ministère de l'Éducation du Québec, avril 1971.
- Ébel, R.L, *Essentials of Educational Measurement*, Englewood Cliffs, N.J. : Prentice-Hall, 1972.
- Gardner, E.F. Bias, *Measurement in Education*, 9, 1978, p. 1-4.
- Guion, R., Content validity — the source of my discontent, *Applied Psychological Measurement*, 1, 1977, p. 1-10.
- Gunnarsson, B, A look at the content similarities between intelligence, achievement, personality, and language tests. In J.W. Oller et K. Perkins, Ed, *Language in Education : testing the tests*, Rowley, Mass. : Newbury House, 1978, ch. 2.
- Hill, J.E., *The Educational Sciences*, Bloomfield Hills : Oakland Community College, 1974.
- Landry, M., *Étude comparative de deux types de rendement en mathématiques (oral et écrit) en fonction d'habiletés de lecture*, Mémoire de maîtrise en éducation, Université du Québec à Montréal, août 1979.
- Morissette, D., *Les examens de rendement scolaire. (Comment les préparer et comment les administrer)*, Québec : Les Presses de l'Université Laval, 1979.
- Oller, J.W. Jr. et Perkins, K., *Language in Education : testing the test*, Rowley, Mass. : Newbury House, 1978.
- Séguin, S.P., *Évaluation, processus décisionnel et rôle des mesures*, Conférence prononcée dans le cadre de la Semaine Pédagogique Nationale, Québec, mai 1979.
- Séguin, S.P. et Bélair-Pepin, L., Rapport final de recherche sur un curriculum formulé en objectifs opératoires pour l'apprentissage des sciences au secondaire, *GREC : Bulletin No 12*. Groupe de Recherche en Évaluation des Curriculum, Département des sciences de l'éducation, Université du Québec à Montréal, juin 1978.
- Streiff, V., Relationships among oral and written cloze scores and achievement test scores in a bilingual setting, in J.W. Oller et K. Perkins, ed., *Language in Education : testing the tests*, Rowley, Mass. : Newbury House, 1978, ch. 4.
- Stump, T.A., Cloze and dictation tasks as predictors of intelligence and achievement scores, in J.W. Oller et K. Perkins, ed. *Language in Education : testing the tests*, Rowley, Mass. : Newbury House, 1978, ch. 3.
- Standards for Educational and Psychological Tests*. Washington, D.C. : American Psychological Association, 1974.