

Identification des compétences à l'école élémentaire : une approche empirique à partir des évaluations institutionnelles

Sophie Morlaix et Bruno Suchaut

Volume 30, numéro 2, 2007

URI : <https://id.erudit.org/iderudit/1085883ar>

DOI : <https://doi.org/10.7202/1085883ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Morlaix, S. & Suchaut, B. (2007). Identification des compétences à l'école élémentaire : une approche empirique à partir des évaluations institutionnelles. *Mesure et évaluation en éducation*, 30(2), 1–22. <https://doi.org/10.7202/1085883ar>

Résumé de l'article

Cet article propose une analyse des résultats des élèves aux évaluations nationales, l'objectif étant de mieux comprendre comment les acquisitions se structurent à l'école primaire. L'intérêt de cette démarche réside principalement dans l'identification de compétences dont la maîtrise est essentielle à la réussite scolaire. L'approche choisie pour traiter la question est résolument empirique et se base sur l'examen des relations statistiques entre les items. Elle permet de répondre à plusieurs questions susceptibles d'intéresser les chercheurs en éducation, mais aussi les acteurs engagés, à différents degrés, dans le pilotage et la gestion pédagogique de l'enseignement primaire. Les évaluations nationales sont en effet surtout exploitées dans une logique diagnostique alors que notre perspective permet de réfléchir sur des prescriptions plus larges en matière de politique éducative, touchant aussi bien la définition des curricula que leurs modalités d'application.

Identification des compétences à l'école élémentaire: une approche empirique à partir des évaluations institutionnelles

Sophie Morlaix

Bruno Suchaut

Université de Bourgogne et IREDU CNRS

MOTS CLÉS: Acquisitions, compétences, items, école primaire, évaluation

Cet article propose une analyse des résultats des élèves aux évaluations nationales, l'objectif étant de mieux comprendre comment les acquisitions se structurent à l'école primaire. L'intérêt de cette démarche réside principalement dans l'identification de compétences dont la maîtrise est essentielle à la réussite scolaire. L'approche choisie pour traiter la question est résolument empirique et se base sur l'examen des relations statistiques entre les items. Elle permet de répondre à plusieurs questions susceptibles d'intéresser les chercheurs en éducation, mais aussi les acteurs engagés, à différents degrés, dans le pilotage et la gestion pédagogique de l'enseignement primaire. Les évaluations nationales sont en effet surtout exploitées dans une logique diagnostique alors que notre perspective permet de réfléchir sur des prescriptions plus larges en matière de politique éducative, touchant aussi bien la définition des curricula que leurs modalités d'application.

KEY WORDS: Acquisitions, skills, items, primary school, evaluation

This paper presents an analysis of students' results at national evaluations. This work aims at understanding how academic acquisitions are built in primary school. The main contribution is an identification of skills, very important in school achievement in the cycle III of the French educational system. This empirical approach is based on data and results in different items. The results lead to some issues on educational policy, mainly about the development of some school activities in relation with curricula.

Note des auteurs – Toute correspondance peut être adressée comme suit: Sophie (Aubriet) Morlaix, IREDU/CNRS, Institut de recherche en éducation-Sociologie et économie de l'éducation, BP 26513, 21065 Dijon Cedex, ou par courriel à l'adresse suivante: [sophie.morlaix@u-bourgogne.fr], Bruno Suchaut [bruno.suchaut@u-bourgogne.fr].

PALAVRAS-CHAVE: Aquisições, competências, itens, escola elementar, avaliação

Este artigo propõe uma análise dos resultados dos alunos nas avaliações nacionais, tendo como objectivo compreender melhor como se estruturam as aquisições na escola elementar. O interesse deste procedimento reside, principalmente, na identificação de competências essenciais ao êxito escolar. A abordagem escolhida para tratar a questão é empírica e baseia-se na análise das relações estatísticas entre os itens. Ela permite responder a várias questões susceptíveis de interessar aos investigadores em educação, mas também aos actores envolvidos, a diferentes níveis, na pilotagem e na gestão pedagógica do ensino obrigatório. As avaliações nacionais são exploradas numa lógica diagnóstica enquanto a nossa perspectiva permite reflectir sobre prescrições mais amplas em matéria de política educativa, abordando quer a definição dos curricula, quer as suas modalidades de aplicação.

Depuis 1989, le système éducatif français s'est doté d'outils d'évaluation des élèves à différents niveaux de la scolarité primaire et secondaire¹. Distribuées et utilisées à l'échelon national, ces évaluations de masse ont, d'une part, contribué de manière progressive à la diffusion de la culture de l'évaluation à l'ensemble de la communauté éducative et, d'autre part, permis aux enseignants d'approfondir leur réflexion pédagogique (Joutard & Thélot, 1999; Ministère de l'Éducation nationale, 2000). L'objectif initial de ces évaluations nationales était de «permettre l'observation des compétences et apprécier les réussites et les difficultés éventuelles de chaque élève considéré individuellement, à un moment précis de la scolarité. Elles fournissent aux enseignants des repères exploitables quant aux types d'erreurs fréquemment produites par les élèves au cours de leur apprentissage [...]»². Même si la dimension diagnostique reste aujourd'hui prioritaire, ces évaluations peuvent être utilisées comme des instruments de pilotage car elles rendent objectivement compte du degré d'acquisition des élèves à plusieurs paliers : national, régional, local.

Pour le chercheur en éducation, les évaluations nationales présentent également un grand intérêt dans le sens où celles-ci balayent un vaste ensemble de compétences et de connaissances alors que les tests couramment utilisés dans les recherches (françaises notamment) sont beaucoup plus restrictifs sur ce plan. Sur la base d'une recherche récente (Morlaix & Suchaut, 2006³), nous essaierons, dans cet article, de montrer en quoi les outils d'évaluation peuvent livrer une information capitale sur les apprentissages réalisés par les élèves, à condition de ne pas se limiter à la seule approche institutionnelle et à la classification des compétences établies par les concepteurs des épreuves.

Notre démonstration se base sur une exploitation des résultats aux épreuves nationales. Des données ont été collectées sur un échantillon d'environ 700 élèves à différents moments de la scolarité. Ces données concernent les résultats détaillés d'une même cohorte d'élèves aux évaluations nationales de 1999, pour le niveau du CE2 (troisième année de l'école élémentaire), et de 2002 pour la classe de sixième (première année de collège); les élèves ont également été testés en fin de cinquième (deuxième année de collège) en 2003. Nous disposons aussi d'autres données sociodémographiques et scolaires pour tous les élèves. Nous n'aborderons dans cet article que la première partie de la recherche évoquée précédemment, à savoir l'analyse des résultats des élèves aux épreuves de CE2⁴.

En 1999, dans les évaluations du ministère de l'Éducation nationale de CE2, sept champs de compétences sont préalablement définis par les concepteurs des épreuves : compréhension (41 items), outils de la langue (40 items), production d'écrits (10 items) pour le français; travaux géométriques (17 items), mesure (32 items), travaux numériques (32 items), et résolution de problèmes (9 items) pour les mathématiques. Les 91 items de français sont répartis dans 18 exercices, tandis que les 80 items de mathématiques sont issus de 27 exercices. Plusieurs séquences (quatre en mathématiques et quatre en français) sont prévues pour faire passer aux élèves l'ensemble de ces épreuves. Les cahiers d'évaluations (un en français et un en mathématiques) remplis par les élèves sont alors corrigés par les enseignants, grâce aux consignes de correction nationales.

Au total, les évaluations nationales de CE2 comportent 171 items, regroupés en 15 compétences de français et 27 de mathématiques; chacune de ces compétences correspond, dans la grande majorité des cas, à un exercice du cahier d'évaluation. Ces exercices reflètent les programmes scolaires officiels en vigueur et constituent à ce titre une référence commune pour l'ensemble des enseignants et des élèves. Sur la base des simples taux de réussite aux items, il est possible d'avoir une vision du niveau d'acquisition des élèves à un moment donné de leur scolarité et de relever les lacunes de chacun en vue d'y remédier au cours de l'année scolaire. Outre cette perspective diagnostique et pédagogique, les résultats des évaluations fournissent des indicateurs objectifs de la réussite des élèves. Le Ministère insiste toutefois sur l'utilisation inadaptée qui pourrait être faite des résultats des évaluations, ceux-ci ne pouvant, en aucun cas, être considérés comme des normes à atteindre, mais seulement comme un diagnostic, en début de cycle⁵, des réussites et des difficultés des

élèves. Il est également clairement énoncé dans le discours officiel que les épreuves d'évaluation ne rendent que partiellement compte des compétences et des connaissances des élèves puisque la passation est collective et limitée à des exercices écrits. La prudence est aussi conseillée quant à l'interprétation de l'évolution des taux de réussite au cours du temps, les épreuves n'étant pas, la plupart du temps, identiques d'une année sur l'autre (Morlaix & Suchaut, 2006). Au-delà de ces limites, l'exploitation des évaluations nationales pose de réels problèmes méthodologiques quand on utilise la catégorisation des compétences proposée par les concepteurs de ces outils. Deux points essentiels méritent d'être soulignés, le premier concerne l'échelle de mesure utilisée, le second la définition des différentes compétences.

L'échelle utilisée pour apprécier la réussite des élèves est très variable selon le nombre d'items retenus pour mesurer chacune des compétences. Par exemple, la compétence «*comprendre un texte...*» est évaluée par 12 items alors que la compétence «*construire une figure simple sur un quadrillage...*» ne concerne qu'un seul item. Selon le cas, la graduation de la réussite est donc très fluctuante d'une compétence à l'autre, allant de la simple dichotomie «échec/réussite» à une échelle en 12 scores. Une autre difficulté est relative au seuil de réussite : à partir de quel score peut-on considérer que l'élève maîtrise la compétence ? On peut bien sûr fixer un seuil de réussite commun à toutes les compétences en adoptant les principes en usage (75 % est généralement le seuil retenu), mais les dissymétries entre les différentes échelles rendent cette solution trop imparfaite, notamment pour les compétences qui comportent très peu d'items.

Indépendamment des problèmes liés à l'échelle de mesure, les évaluations nationales peuvent donner lieu à des interrogations encore plus profondes concernant la définition même des compétences évaluées, ces compétences étant définies par les commissions chargées de l'élaboration des épreuves. Quand on teste la pertinence des regroupements d'items qui composent les compétences, deux observations majeures ressortent. En premier lieu, il apparaît que des items provenant de compétences différentes, voire de disciplines différentes, entretiennent des liens statistiques très forts. On peut ainsi imaginer que les items figurant dans les épreuves puissent s'associer différemment pour rendre compte des compétences des élèves. En second lieu, les items censés rendre compte d'une même compétence peuvent n'être que très faiblement corrélés (Morlaix & Suchaut, 2006). C'est, par exemple, le cas de la compétence «*comprendre et savoir appliquer les consignes courantes du travail scolaire*», pour laquelle la plupart des corrélations entre les cinq items la composant ne sont pas significatives.

De nombreux autres exemples pourraient être évoqués pour montrer qu'une définition *a priori* des compétences n'est pas adéquate. L'approche institutionnelle relève d'un certain arbitraire dans la définition des compétences et il n'est pas certain que chaque item mesure réellement la compétence visée. En outre, cette approche peut être considérée comme statique dans la mesure où elle se limite à une catégorisation par domaine d'acquisition et par discipline. En cela, elle ne permet pas de prendre en compte la logique d'interdépendance des acquis des élèves.

Toutes ces remarques soulignent la difficulté à saisir concrètement la notion de compétence dans le domaine de l'éducation (Crahay, 2006). Même si, d'un point de vue théorique, les auteurs s'accordent sur une définition (une compétence renverrait à un ensemble intégré de connaissances susceptibles d'être mobilisées pour accomplir des tâches), la question de la mesure demeure. Pour progresser dans ce domaine, nous nous proposons de réperer, non plus *a priori*, mais *a posteriori* et de manière empirique, les diverses compétences mises en œuvre par les élèves à partir des situations d'évaluation figurant dans les épreuves nationales. Notre travail s'inspire globalement des procédures utilisées pour les validations empiriques *a posteriori* des épreuves d'évaluation (De Ketele & Gérard, 2005). La démarche méthodologique que nous avons adoptée permet d'appréhender les apprentissages des élèves de façon plus précise et plus dynamique en centrant l'analyse sur les items et non sur les compétences définies préalablement dans les épreuves. Même si cette approche comporte elle aussi des limites, notamment quant au degré de difficulté des items (Demeuse & Henry, 2004), elle présente des avantages certains.

L'objectif principal de notre travail est donc de mieux comprendre comment les acquisitions des élèves s'articulent en précisant des structures capables de dresser une cartographie des compétences acquises par les élèves à un moment donné de la scolarité primaire, en l'occurrence, à l'entrée au cycle III. Le terme de compétence intègre ici, dans sa définition, les différentes capacités, aptitudes, connaissances et attitudes qui mobilisent chez les individus des ressources cognitives pour réaliser une tâche.

Dans un premier temps, nous soulignerons l'intérêt de l'approche centrée sur les items, puis nous présenterons la démarche statistique suivie pour identifier empiriquement les compétences. Les principaux résultats des analyses, à savoir la structure relationnelle et hiérarchique des compétences, seront ensuite exposés. Enfin, les conséquences possibles de notre travail en matière de politique éducative seront évoquées.

Une approche basée sur les relations statistiques entre items

L'approche centrée sur les items a un double objectif : tester statistiquement la pertinence de la nomenclature des compétences établie par les concepteurs des évaluations nationales et identifier de nouvelles compétences à partir d'associations d'items. D'un point de vue concret, cette approche permet d'utiliser une échelle de mesure commune puisque tous les items présentent le même barème de cotation : 0 pour une réponse erronée, 1 pour la réponse attendue. La question du seuil de réussite mentionnée auparavant ne se pose plus, puisque ce seuil est défini objectivement par la réussite ou l'échec, sans possibilité de situations intermédiaires⁶.

Le point de départ de notre démarche est la production d'une matrice de corrélations qui intègre tous les items de l'épreuve de CE2 (171 items de français et de mathématiques). Compte tenu du nombre important de corrélations (14 535 coefficients de corrélation), nous avons décidé de ne retenir que celles supérieures à +0,20, soit 317⁷ ; on notera déjà que cette sélection de corrélations intègre la totalité des items de l'évaluation. Sans grande surprise, on remarque qu'une bonne partie des corrélations (57% d'entre elles) se rapporte à des items appartenant à un même exercice et 86%, à la même discipline. Par complémentarité, on observe aussi qu'un nombre non négligeable de corrélations concerne des items provenant d'exercices différents : cela confirme l'observation précédente selon laquelle certaines compétences seraient mesurées de façon imparfaite.

Afin d'organiser l'analyse des relations entre items, une phase préparatoire consiste à étudier individuellement chaque corrélation et à dresser ainsi une cartographie de l'ensemble des situations présentes. Le principe de cette étape préalable est de relever des blocs de relations au sein desquelles on retrouve le plus souvent les mêmes items. Cette procédure revêt, de fait, un caractère systématique puisque, pour chaque item, on établit tous les autres items qui lui sont associés dans les corrélations. Au terme de cette phase, on aboutit à des groupements d'items fortement corrélés entre eux ; certains de ces groupements font intervenir un nombre important de corrélations alors qu'à l'opposé, d'autres ne contiennent que deux ou trois corrélations, celles-ci n'ayant pu être rattachées à d'autres regroupements. Au total, 29 blocs d'items ont été déterminés. À titre d'illustration, le tableau 1 présente l'un de ces blocs d'items (il s'agit de huit items reliés les uns aux autres par neuf corrélations bivariées).

Tableau 1
Exemple d'un regroupement d'items sur la base des corrélations⁸
(évaluation CE2, 1999)

<i>Corrélations</i>	<i>Description succincte des items</i>
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F62	Dictée préparée d'une phrase : bonne orthographe du mot « pendant »
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F60	Respect de la ponctuation dans une production écrite
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F64	Dictée préparée d'une phrase : bonne orthographe du mot « garçons »
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F66	Dictée préparée d'une phrase : bonne orthographe du mot « filles »
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F63	Dictée préparée d'une phrase : bonne orthographe du mot « récréation »
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F12	Repérer des types d'écrits à partir d'extraits de livres différents
F67	Dictée préparée d'une phrase : bonne orthographe du mot « jouent »
F16	Répondre à des questions concernant un texte
F64	Dictée préparée d'une phrase : bonne orthographe du mot « garçons »
F62	Dictée préparée d'une phrase : bonne orthographe du mot « pendant »
F63	Dictée préparée d'une phrase : bonne orthographe du mot « récréation »
F62	Dictée préparée d'une phrase : bonne orthographe du mot « pendant »

Parmi ces huit items de français, cinq items appartiennent au même exercice (60, 62, 63, 64, 66, 67) et sont censés mesurer la même compétence «*écrire sous la dictée des mots courants, de petites phrases ou de petits textes*». Dans cet exercice, il est demandé aux élèves d'écrire la phrase suivante, dictée par l'enseignant : «Pendant la récréation, les garçons et les filles jouent aux billes.» Cette phrase a été auparavant écrite au tableau, les élèves étant invités à mémoriser l'orthographe des mots. La phrase a été ensuite effacée et un exercice différent a été proposé aux élèves ; ce n'est qu'après cet exercice que la phrase a été l'objet d'une dictée. La réussite à la dictée dépend de plusieurs éléments : de compétences orthographiques, de la capacité à se remémorer la graphie des mots préalablement observés et, dans une certaine mesure, d'une capacité d'attention. L'item 60 est supposé rendre compte de la compétence «*copier un mot, une phrase, un texte...*». Cet item fait partie d'un exercice dans lequel il est demandé aux élèves de recopier des phrases faisant partie d'un court texte présenté sous la forme d'une constellation (les phrases sont écrites dans des bulles et reliées par une flèche à un dessin qui représente

le thème du texte) et de les réorganiser sous une forme plus conventionnelle (un texte organisé en paragraphes). L'item 60 n'évalue qu'un aspect de la production des élèves : le respect de la ponctuation (majuscules et points).

L'item 12 fait partie d'un exercice dans lequel on demande aux élèves de repérer des types d'écrits à partir d'un extrait de livres différents (poésie, dictionnaire, livre de recettes de cuisine, etc.). La compétence visée est « *distinguer des textes différents en utilisant des indices extérieurs au texte* ». Enfin, l'item 16 vise la compétence « *comprendre un texte et montrer qu'on l'a compris* ». Un texte est proposé aux élèves et ceux-ci doivent répondre à des questions concernant ce texte. Dans l'item 16, les élèves doivent entourer trois bonnes réponses parmi six propositions.

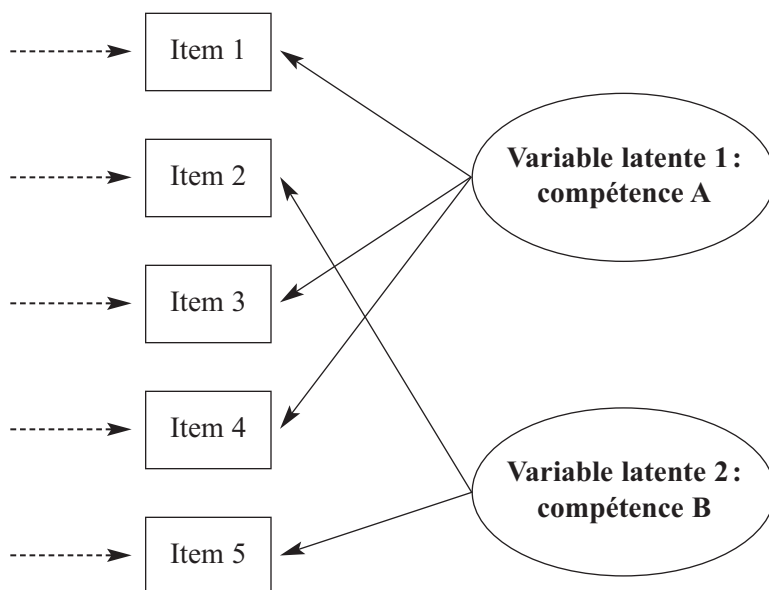
Selon notre logique d'analyse, ce regroupement d'items fortement corrélés serait censé rendre compte d'une ou de plusieurs compétences. On rappellera que, dans les évaluations nationales, ces huit items correspondent à quatre compétences distinctes. L'étape suivante consiste à analyser les liaisons statistiques à l'intérieur de chaque regroupement d'items de façon à mettre à jour des variables latentes pouvant être interprétées comme des compétences, des aptitudes ou des capacités mobilisées par les élèves dans les évaluations nationales. Pour cela, des modèles de mesure ont été estimés à l'aide du logiciel LISREL⁹.

Une analyse en variables latentes

Les modèles de mesure permettent de rendre explicite une dimension latente des phénomènes observés en postulant l'existence de variables inobservées (ou inobservables) qui rendent compte des relations entre variables observées, en l'occurrence ici les résultats aux items. La démarche avec LISREL implique plusieurs étapes (Aish-Van Vaerenbergh, 1997), dans la mesure où le logiciel permet l'estimation de deux types de modèles. Le premier se rapporte aux modèles de mesure qui établissent les relations entre les variables latentes (compétences supposées) et leurs indicateurs (les items), le second concerne le modèle structurel qui détermine les liaisons entre les différentes variables latentes. L'analyse réalisée ici sur les items de CE2 ne concerne que les modèles de mesure.

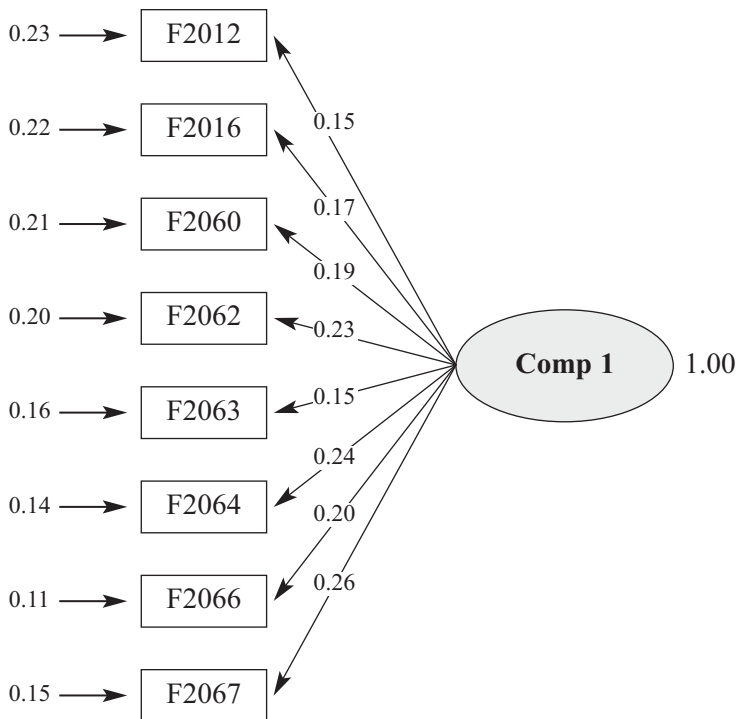
Il est d'usage de représenter le modèle de mesure par un diagramme dans lequel les variables latentes sont symbolisées par des ellipses et les variables observées (ou indicateurs de ces variables latentes) sont matérialisées par des rectangles (graphique 1). Les flèches en traits pleins matérialisent l'intensité des relations qui lient chacun des indicateurs (items) à la variable latente

(compétence), un coefficient de régression (et son degré de significativité) pour chaque indicateur fournit une indication sur la validité du modèle¹⁰ (Morlaix, 2002). LISREL donne également une mesure de fiabilité du modèle grâce à l'indication de l'erreur de mesure sur chacun des indicateurs (flèches en pointillés sur la figure). Il est donc possible de savoir si les indicateurs retenus sont «robustes» pour mesurer la variable latente.



Graphique 1. *Modèle de mesure entre deux variables latentes et leurs indicateurs*

Des modèles de mesure ont été estimés pour chacun des regroupements d'items; à titre d'exemple, nous allons présenter l'analyse qui concerne le regroupement d'items évoqué précédemment. Une première étape consiste à définir une variable latente qui rend compte des liens entre les items (graphique 2).



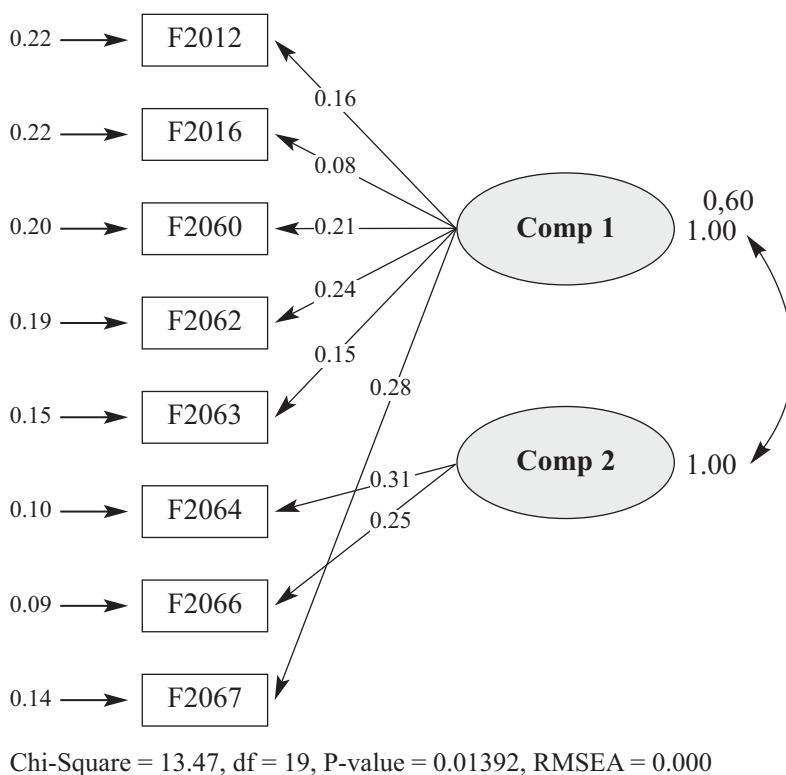
Chi-Square = 70.16, df = 20, P-value = 0.00000, RMSEA = 0.066

Graphique 2. *Modèle de mesure initial*

On peut lire sur les flèches les coefficients des régressions entre chaque item et la variable latente (comp1). Les valeurs des coefficients nous renseignent sur la capacité que peut avoir l’item à rendre compte de la variable latente, celle-ci étant, rappelons-le, censée traduire une compétence. C’est pour l’item 67 que la variable latente est la plus prédictive (coefficient de +0,26). D’autres statistiques fournies par LISREL indiquent également que l’item 67 explique 31 % de la variance de la variable latente. Les statistiques fournies par LISREL (Khi^2 , P.Value et les résidus) indiquent que ce premier modèle de mesure peut être amélioré et le logiciel suggère qu’une nouvelle variable latente pourrait être introduite dans l’analyse en isolant les items 64 et 66.

Sur la base de cette nouvelle information, un modèle faisant intervenir deux variables latentes est estimé. La première variable est déterminée par les items 12, 16, 60, 62, 63 et 67 («comp1» sur le graphique), la seconde par les items 64 et 66 («comp2» sur le graphique). Les indicateurs statistiques du

nouveau modèle deviennent beaucoup plus fiables (baisse sensible du Khi^2 , augmentation de la P.Value, baisse du RMSEA) et le logiciel ne propose alors aucune autre variante pour améliorer la qualité statistique de la représentation. Le graphique ci-après traduit les nouvelles relations entre les items et les deux variables latentes (graphique 3).



Graphique 3. *Modèle de mesure amélioré*

On voit aussi apparaître sur le schéma la corrélation entre les deux variables latentes, qui affiche une valeur de +0,60: cela témoigne de la proximité des deux compétences. Toutes les relations matérialisées sur le schéma sont en outre significatives. Cette nouvelle configuration demande à être interprétée. Il semble donc que les items 64 et 66 mesurent une compétence spécifique qui rend compte des aptitudes des élèves à orthographier correctement des noms écrits au pluriel («garçons» et «filles»).

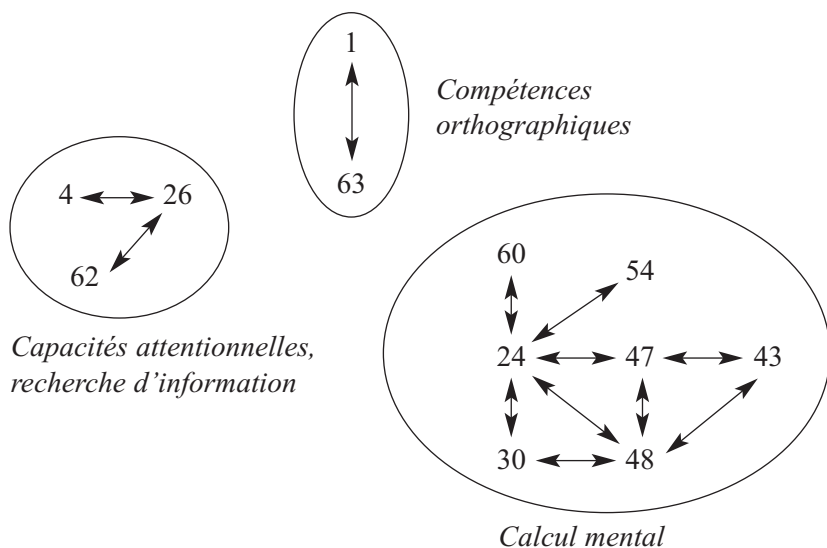
Les six autres items (12, 16, 60, 62, 63 et 67) mesureraient, pour leur part, une compétence différente. Il est difficile d'interpréter rigoureusement ces relations sans avoir recours à des éléments théoriques. En effet, les résultats des élèves aux items de ce premier regroupement semblent dépendre de plusieurs aptitudes : orthographiques, mémorielles, attentionnelles, etc.

Les modèles de mesure estimés sur la totalité des regroupements d'items ont donné lieu à l'identification de 63 variables latentes (Morlaix & Suchaut, 2006). L'analyse en variables latentes a donc conduit à une recombinaison des compétences des élèves à l'entrée au CE2. Parmi les 63 variables mises à jour, 27 d'entre elles (soit 43 %) correspondent, souvent de façon partielle, à des regroupements d'items déjà présents dans les évaluations nationales. La correspondance entre les compétences des épreuves et les variables latentes est néanmoins très imparfaite puisque seules cinq variables correspondent exactement à des compétences figurant dans les évaluations nationales (il s'agit uniquement d'exercices de mathématiques).

Structure relationnelle et hiérarchique des compétences

Afin de mettre à jour la structure des acquisitions des élèves, il est à présent indispensable d'étudier les relations statistiques entre les variables latentes. Cela permettra notamment de répondre à une question centrale : quelles compétences rendent le mieux compte des acquisitions globales des élèves ? Pour répondre à cette question, nous avons, dans une première étape, analysé la matrice de corrélations entre les différentes variables latentes. Les coefficients de corrélation présentent des valeurs allant de 0 à +0,80 (la moyenne sur l'ensemble des corrélations étant de +0,24). Nous avons sélectionné les corrélations les plus importantes (supérieures à +0,70). Cela a permis de dégager trois grands groupes de compétences qui sont représentés sur le graphique 4 (les chiffres correspondent aux numéros associés aux variables latentes). Un premier ensemble est composé des variables «comp1» et «comp63». Il s'agit ici clairement de compétences orthographiques qui sont mesurées avec ces deux variables. En effet, les items rendant compte de ces variables latentes concernent, presque en totalité, deux exercices de dictée. Le deuxième ensemble regroupe les variables «comp4», «comp26» et «comp62». Cet ensemble d'apparence disparate prend sens quand on l'examine sous l'angle de la psychologie cognitive¹¹. Autant les items de français que ceux de mathématiques présents dans ce regroupement évaluent la capacité que peuvent avoir les élèves à rechercher de l'information plus ou moins complexe à

partir de supports divers (textes, mots, calendriers, emplois du temps, plans, énoncés de problème). Ce sont donc les capacités attentionnelles des élèves qui sont mises à contribution pour la maîtrise de cette compétence globale. Le troisième ensemble regroupe sept variables latentes. Le lien commun entre ces variables est également clair puisque les items de calcul mental interviennent systématiquement pour chacune d'entre elles.



Graphique 4. *Blocs de compétences des évaluations CE2*

Les acquisitions des élèves à l'entrée au CE2 s'organisent principalement autour de ces trois compétences qui ne sont pas de même nature. Si l'acquisition des compétences orthographiques dépend principalement d'un enseignement systématique, les deux autres compétences sont davantage associées à des processus plus complexes qui interviennent de façon transversale dans de nombreuses situations d'apprentissage. On devrait s'attendre à ce que ces compétences majeures contribuent fortement à l'explication des différences de réussite entre élèves à l'entrée au CE2. Pour vérifier cela, une régression « pas à pas » a été estimée avec, comme variable dépendante, le score global moyen de CE2 et, comme variables explicatives, l'ensemble des compétences mises à jour précédemment. Le tableau suivant présente les résultats en ne conservant que les 12 compétences les plus prédictives. Celles-ci expliquent, à elles seules, plus de 92 % de la variance du score global, soit presque sa totalité.

On constate que les trois compétences les plus prédictives («comp48», «comp4», «comp63») appartiennent chacune à un des groupes identifiés auparavant. Ces trois compétences expliquent à elles seules 82% de la variance du score global. Cette analyse de la prédictivité montre que certaines compétences sont au cœur des acquisitions des élèves à l'entrée au CE2. Les habiletés en calcul mental, la capacité à retrouver rapidement des données dans des supports variés, la maîtrise de l'orthographe, structurent ainsi fortement les résultats des élèves au début du cycle III.

Tableau 2
*Modèle analysant la prédictivité des variables latentes
sur le score global de CE2 (régression pas à pas)*

<i>Ordre d'introduction</i>	<i>Variabes</i>	<i>Coefficients standardisés</i>	<i>t de Student</i>
1	Comp48	+0,23	12,3 ***
2	Comp4	+0,24	14,7 ***
3	Comp63	+0,17	12,4 ***
4	Comp58	+0,13	10,2 ***
5	Comp13	+0,12	10,0 ***
6	Comp17	+0,11	8,9 ***
7	Comp60	+0,13	7,2 ***
8	Comp41	+0,08	6,6 ***
9	Comp16	+0,08	6,8 ***
10	Comp34	+0,08	7,2 ***
11	Comp15	+0,08	6,8 ***
12	Comp3	+0,08	6,5 ***
	Constante	68,76	510,02
	R ²	0,92	

*** : significatif au seuil de 1%

Une autre question est d'examiner comment ces compétences les plus prédictives des évaluations nationales se hiérarchisent entre elles¹². Le tableau 3 expose les relations entre la réussite et l'échec aux 12 compétences les plus prédictives du score global de CE2. Les compétences sont présentées de façon hiérarchisée dans le tableau (de «comp15» à «comp3»). La réussite à la première compétence «comp15» est donc censée être une condition à la réussite de la seconde («comp58»), qui elle-même détermine la réussite à la suivante («comp41») etc., jusqu'à la dernière compétence («comp3») qui ne peut être acquise que si toutes les autres compétences sont maîtrisées. Dans chaque

case, le nombre indiqué représente le pourcentage d'élèves (sur la totalité de l'échantillon) qui réussit une compétence tout en ayant échoué à une autre (il s'agit de résultats issus de tableaux croisés). Ainsi 4% des élèves réussissent «comp58» tout en échouant à «comp15»; de même, 3,3% des élèves réussissent «comp41» et échouent à «comp15». Si la structure hiérarchique des compétences se vérifie, on devrait donc s'attendre à ce que les chiffres du tableau présentent des valeurs faibles. Cela est effectivement le cas, même si certaines valeurs s'écartent de la règle générale. Ainsi, si la maîtrise de «comp15» semble bien être une condition indispensable à la réussite à toutes les autres compétences (la première ligne du tableau présente des valeurs très faibles), la relation entre «comp13» et «comp16» ne semble pas vérifiée dans la mesure où plus de 20% des élèves réussissent «comp13» tout en échouant à «comp16». Globalement, la structure de dépendance entre compétences se vérifie dans la majorité des cas; les cases grisées dans le tableau représentent ces situations (pourcentages d'élèves inférieurs à 10% ou voisins de ce pourcentage).

Tableau 3

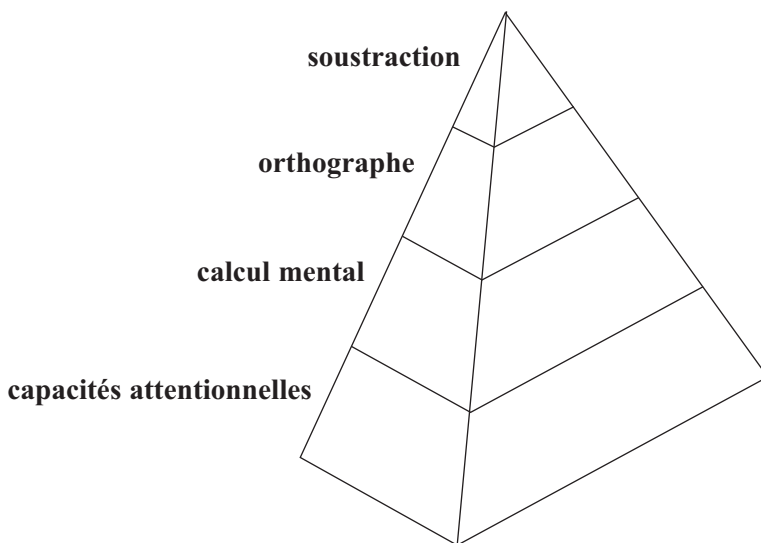
Relations entre l'échec et la réussite aux compétences de CE2 (% d'élèves)

RÉUSSITE												
ÉCHEC	comp 15	comp 58	comp 41	comp 17	comp 16	comp 13	comp 34	comp 4	comp 48	comp 60	comp 63	comp 3
comp15		4,0	3,3	2,1	3,0	2,4	2,8	0,0	0,4	0,1	0,1	0,6
comp58			8,5	6,7	7,9	7,6	8,0	1,2	1,3	0,7	0,7	1,2
comp41				14,6	14,6	14,0	13,7	2,8	3,1	1,6	1,3	1,6
comp17					19,5	17,7	16,1	4,5	4,9	3,6	2,5	3,1
comp16						20,6	16,7	6,7	6,7	4,3	4,8	3,1
comp13							16,5	4,5	6,1	4,2	2,2	3,3
comp34								6,9	7,2	4,6	5,0	4,9
comp4									11,2	7,9	6,7	8,9
comp48										5,8	7,5	6,1
comp60											8,6	7,5
comp63												9,8

Les compétences semblent donc assez bien hiérarchisées dans l'ensemble; la compétence qui se trouve au sommet de cette hiérarchie correspond à la maîtrise de la soustraction (variable «comp3»). Ceci signifie que l'acquisition de la technique opératoire de la soustraction est un processus qui nécessite de

la part des élèves diverses capacités et habiletés préalables. En effet, très peu d'élèves réussissent dans cette dimension des acquisitions sans maîtriser toutes les autres compétences qui figurent dans le tableau 3. La seconde compétence dont la maîtrise dépend de nombreuses autres est de nature orthographique («comp63»); viennent ensuite deux compétences dans lesquelles le calcul mental intervient majoritairement, soit directement («comp48»), soit indirectement par le biais d'autres compétences («comp60»). La compétence qui se situe au plus bas niveau de cette structure hiérarchique correspond à des items évaluant la connaissance des unités de mesure. En fait, c'est davantage la compréhension de consignes simples qui semble être la compétence véritablement visée dans cet exercice. Le tableau en annexe fournit la liste de l'ensemble des compétences du tableau en indiquant les items qui s'y réfèrent. Le graphique 5 symbolise quant à lui la structure pyramidale des dernières compétences du même tableau.

Cette structure pyramidale des compétences peut fournir des indications didactiques et pédagogiques pour l'enseignement au cycle II, en amont du CE2. Ces indications concernent principalement les contenus d'enseignement et leur programmation dans le temps. Ils s'appuient sur le double constat suivant établi à l'entrée du CE2: certaines compétences sont difficilement accessibles à l'ensemble des élèves, certaines compétences sont essentielles à l'acquisition d'autres compétences.



Graphique 5. *Structure pyramidale des compétences des élèves au CE2*

Après ces deux constats, on peut déjà suggérer que la définition des programmes scolaires mette l'accent de façon explicite sur les compétences définies comme essentielles précédemment. Cela nécessite aussi que des situations d'enseignement concrètes puissent être envisagées pour l'acquisition de ces compétences, ce qui n'est pas immédiat quand elles sont transversales. Il est ensuite utile de s'interroger sur la planification de l'enseignement de ces compétences en prenant en compte la structure hiérarchique d'apprentissage telle qu'elle apparaît sur le graphique précédent, c'est-à-dire en respectant le fait que certains apprentissages ne peuvent se réaliser que si certains autres sont déjà maîtrisés.

Conclusion

L'objectif de cet article était d'étudier la structure des acquisitions des élèves à l'entrée en CE2 et d'identifier, par une approche empirique, les compétences principales. Un premier constat qui se dégage des résultats est l'aspect transversal des acquisitions scolaires. La vision institutionnelle des évaluations nationales masque cet aspect par le regroupement, parfois artificiel, de certains items au sein d'un même exercice alors que l'approche statistique permet d'identifier des compétences de nature variée. De façon complémentaire, des items provenant d'exercices variés, voire de disciplines différentes ont été identifiés comme mesurant une seule et même compétence. Ce premier constat peut avoir des implications évidentes quant à l'utilisation des évaluations nationales sur le plan pédagogique, dans une perspective diagnostique. Cela permet en effet de mieux cerner les processus d'apprentissage des élèves qui sont mobilisés lors de la réalisation d'un exercice et donc de fournir des éléments pour organiser des situations pédagogiques ciblées sur l'acquisition de certaines compétences.

Parmi la soixantaine de compétences identifiées, certaines ont un statut particulier dans le sens où elles rendent compte, à elles seules, de la presque totalité des différences d'acquisition des élèves à l'entrée au cycle III. Ces compétences concernent trois domaines : l'orthographe, l'attention et la recherche d'information, le calcul mental¹³. On rejoint ici un constat réalisé par la DEP sur la base du panel 1997¹⁴ selon lequel l'épreuve d'attention partagée administrée à l'entrée au CP s'est révélée fortement corrélée avec la réussite générale à l'évaluation de CE2, deux années plus tard ; «la capacité d'attention pourrait donc bien constituer un facteur majeur de la réussite» (Ministère de l'Éducation nationale, 2000).

Notre recherche a aussi permis d'approfondir les relations entre les compétences en mettant en évidence une structure hiérarchique (Morlaix & Suchaut, 2006), certaines compétences étant indispensables à l'acquisition d'autres compétences. Sans entrer ici dans le détail de cette structure, signalons que la compétence de CE2 qui se trouve au sommet de la hiérarchie correspond à la maîtrise de la soustraction, qui apparaît être un processus nécessitant de la part des élèves diverses capacités et habiletés préalables. La compétence se situant au plus bas niveau de la structure hiérarchique concerne des exercices qui visent la compréhension de consignes simples (alors que les évaluations nationales ont répertorié ces exercices dans le domaine de la mesure). En prenant en compte la structure hiérarchique d'apprentissage effective des élèves, c'est-à-dire en respectant le fait, trivial à présent, que certains apprentissages ne peuvent se réaliser que si d'autres sont déjà maîtrisés (Bloom, 1979), les résultats évoqués précédemment débouchent naturellement sur un questionnement lié à la planification des activités d'enseignement, et plus largement au contenu des programmes scolaires. Il est aussi essentiel de s'interroger sur le temps que l'on doit consacrer à l'acquisition des différentes compétences, sachant que le temps alloué est très lié aux progrès des élèves et que celui-ci s'inscrit, pour l'enseignant, dans une logique d'arbitrage entre des activités diverses et concurrentielles (Morlaix, 2000; Suchaut, 1996). En outre, certains élèves ont besoin de plus de temps que d'autres pour acquérir certaines notions. Les recommandations pédagogiques doivent donc aussi tenir compte de la diversité des élèves et de leur rythme d'apprentissage spécifique.

NOTES

1. Les évaluations nationales mises en place par la DEP (Direction de l'évaluation et de la prospective) ont concerné à l'origine trois niveaux de la scolarité : CE2 (troisième année de l'enseignement élémentaire), sixième (première année de l'enseignement secondaire) et seconde (cinquième année de l'enseignement secondaire).
2. Note de service 2002-105 du 30.04.2002, B.O. n° 19 du 9 mai 2002.
3. Cette recherche fait suite à un appel d'offre de l'UNSA-Éducation (centre H. Aigueperse) pour l'IRES (Institut de recherches économiques et sociales).
4. Les données analysées ici concernent la classe de CE2. L'échantillon comporte 671 élèves scolarisés dans 31 écoles d'une circonscription de la Côte d'Or. Nous disposons pour tous ces élèves des résultats détaillés aux évaluations nationales de septembre 1999 (résultats item par item) et de données sociodémographiques : mois et année de naissance, sexe, catégorie socioprofessionnelle du père de l'élève.
5. L'enseignement primaire français est organisé en trois cycles. Le cycle I (cycle des apprentissages premiers) concerne l'école maternelle ; le cycle II (cycle des apprentissages fondamentaux) regroupe la grande section de maternelle, le cours préparatoire et le CE1 ; le cycle III (cycle des approfondissements) intègre le CE2, le CM1 et le CM2.

6. On mentionnera que cette dichotomie (échec/réussite) n'est pas adaptée à une interprétation diagnostique des évaluations nationales et que les barèmes de cotation de certains items comportent à l'origine d'autres paliers qui permettent une analyse des erreurs des élèves.
7. Même si ce seuil revêt un caractère arbitraire, l'analyse serait peu pertinente si tous les liens entre items étaient examinés en détail. En effet, dans un grand nombre de cas ces liens sont, soit non significatifs, soit très faibles, et il est alors difficile de commenter et de donner un sens à ces constatations.
8. Ce sont les relations entre items qui figurent dans la colonne «corrélations» du tableau; par exemple «F67 F62» traduit la corrélation entre les items 67 et 62 de français.
9. LISREL : LInear Structural RELationship.
10. LISREL fournit également un certain nombre de statistiques générales rendant compte de la qualité du modèle de mesure. Le χ^2 (khi carré) mesure la distance séparant les deux matrices de corrélations, celle théorique et celle portant sur les corrélations estimées. Cet indicateur du χ^2 doit être le plus petit possible et il doit être aussi proche que possible du nombre de degrés de liberté du modèle. La π value (P.Value) mesure la probabilité d'obtenir la valeur du χ^2 correcte. Le modèle est considéré comme recevable si cette valeur est supérieure à 0,5 et on cherchera en outre à ce qu'elle soit la plus élevée possible. Enfin, l'indice RMSEA (Root Mean Square Error of Approximation) se rapporte à la moyenne des résidus du modèle et à leur significativité. Le modèle est acceptable si la valeur de cet indicateur est inférieure ou égale à 0,05.
11. Des collègues psychologues du LEAD (Laboratoire d'étude de l'apprentissage et du développement - Université de Bourgogne), Pierre Barrouillet et Valérie Camos, ont été associés à cette partie de la recherche. Ils ont notamment contribué à donner une signification à certaines variables latentes identifiées par les modèles LISREL.
12. Une étape préalable à l'analyse est de rendre compte systématiquement de la maîtrise ou de l'échec à une compétence, ce qui suppose une transformation de l'échelle de mesure. On passe alors d'une échelle d'intervalle (qui varie selon la compétence considérée) à une échelle nominale, représentée de façon dichotomique : échec ou réussite. Il se pose alors la question du seuil à partir duquel on va considérer que l'élève a acquis la compétence visée; il n'existe pas de réponse parfaite, mais nous avons considéré qu'un score supérieur à 75% de réussite soit associé à la maîtrise de la compétence.
13. Ces deux derniers éléments pouvant intervenir dans des situations d'apprentissage très variées, qui dépassent largement l'approche disciplinaire.
14. Élèves entrés au CP en 1997.

RÉFÉRENCES

- Aish-Van Vaerenbergh, A.M. (1997). Modèles statistiques et inférences causales : analyse de structures de covariances avec LISREL. In A.M. Aish-Van Vaerenbergh et al., *Faut-il chercher aux causes une raison? L'explication causale en sciences humaines* (pp. 106-130). Librairie philosophique Vrin.
- Bloom, B.S. (1979). *Caractéristiques individuelles et apprentissages scolaires*. Bruxelles : Labor; Paris : Nathan.

- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*, 154, 97-110.
- De Ketele, J.M., & Gerard, F.M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*, 28(3), 1-26.
- Demeuse, M., & Henry, G. (2004). La théorie classique des tests. In M. Demeuse, *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation*. Les Éditions de l'Université de Liège.
- Joutard, P., & Thélot, C. (1999). *Réussir l'école*. Le seuil, 280 p.
- Ministère de l'éducation nationale (Direction de l'enseignement scolaire) (2000). *L'exploitation de l'évaluation nationale en CE2: la lecture. Programme national de pilotage: actes du séminaire national*. CRDP (Centre régional de documentation pédagogique), Académie de Versailles.
- Morlaix, S. (2000). Rechercher une meilleure répartition du temps scolaire en primaire pour favoriser la réussite au collège. *Revue française de pédagogie*, 130, 121-131.
- Morlaix, S. (2002). Intérêts et apports de l'analyse des variables latentes pour le chercheur en sciences sociales: exemple d'application à l'économie de l'éducation. *Orientation scolaire et professionnelle*, 31(1) 117-138.
- Morlaix, S., & Suchaut, B. (2006). *Évolution et structure des compétences des élèves à l'école élémentaire et au collège. Une analyse empirique des évaluations nationales*. Rapport pour l'UNSA-Éducation. 196 p. et annexes.
- Suchaut, B. (1996). La gestion du temps à l'école maternelle et primaire: diversité des pratiques et effets sur les acquisitions des élèves. *L'année de la recherche en sciences de l'éducation*, 123-153.

Annexe

Structure hiérarchique des compétences des élèves au CE2

<i>Variable latente</i>	<i>N° item</i>	<i>Description succincte des items</i>
Comp3	M2056	Effectuer une soustraction
	M2057	Effectuer une soustraction
Niveau 9	M2058	Effectuer une soustraction
	M2077	Résoudre un problème en maîtrisant la technique de la soustraction
Comp63	F2043	Comprendre un message en s'appuyant sur des indices
	F2060	Respect de la ponctuation dans une production écrite
Niveau 8	F2062	Dictée préparée d'une phrase : bonne orthographe d'un mot
	F2063	Dictée préparée d'une phrase : bonne orthographe d'un mot
	F2067	Dictée préparée d'une phrase : bonne orthographe d'un mot
	F2070	Écrire une phrase dictée par segments : bonne orthographe d'un mot
	F2071	Écrire une phrase dictée par segments : bonne orthographe d'un mot
	F2075	Écrire une phrase dictée par segments : bonne orthographe d'un mot
	M2075	Résoudre un problème à une opération
Comp60	F2041	Comprendre un message en s'appuyant sur les indices
	M2005	Tracer un carré en s'aidant d'un quadrillage
Niveau 7	M2018	Se repérer dans la journée
	M2034	Comparer des distances à partir d'un plan
	M2035	Comparer plusieurs grandeurs simultanément entre elles
	M2038	Choisir l'unité de temps qui convient pour compléter un texte
	M2041	Calcul mental
	M2077	Résoudre un problème en maîtrisant la technique de la soustraction
	M2078	Résoudre une situation de partage ou de groupement
Comp48	F2057	Interprétation d'un schéma rôle et localisation de la légende
	M2031	Utiliser le calendrier : inférence pour répondre à la question
Niveau 6	M2034	Comparer des distances à partir d'un plan
	M2035	Comparer plusieurs grandeurs simultanément entre elles
	M2040	Calcul mental
	M2041	Calcul mental
	M2048	Calcul mental
	M2059	Calcul d'un produit
	M2076	Résoudre un problème à une opération
Comp4	F2005	Se repérer dans la journée d'après un emploi du temps
	F2012	Repérer des types d'écrits à partir d'extraits de livres différents
Niveau 5	F2014	Répondre à une question sur un texte

	F2021	Reconstituer la chronologie des événements dans des textes
	F2023	Reconstituer la chronologie des événements dans des textes
	F2044	Identifier certains aspects d'un texte (trouver des indices du féminin)
	M2018	Se repérer dans la journée
	M2029	Utiliser le calendrier
	M2031	Utiliser le calendrier (inférence pour répondre à la question)
	M2037	Choisir l'unité la mieux adaptée à un mesurage
Comp13	F2055	Comprendre l'organisation d'une page imprimée
	F2066	Dictée préparée d'une phrase : bonne orthographe d'un mot
	M2023	Tracer un segment
Comp16	M2004	Utiliser les instruments de dessin pour achever un tracé
	M2006	Compléter par symétrie une figure dessinée sur un quadrillage
	M2007	Compléter par symétrie une figure dessinée sur un quadrillage
	M2008	Compléter par symétrie une figure dessinée sur un quadrillage
Comp17	F2085	Décrire un personnage de fiction en s'aidant d'un autre texte
	F2087	Écrire le corps d'un récit
	F2088	Écrire le corps d'un récit
	F2089	Écrire le corps d'un récit
	F2091	Écrire le corps d'un récit
Comp34	M2010	Se repérer spatialement, à la suite de la description d'un endroit
Niveau 4	M2011	Se repérer spatialement sur un plan
Comp41	M2051	Poser une addition et l'effectuer
Niveau 3	M2052	Poser une addition et l'effectuer
	M2066	Écrire des nombres du plus petit au plus grand
Comp58	F2013	Comprendre un texte et montrer qu'on l'a compris
Niveau 2	F2076	Transformer un texte en appliquant des règles simples (féminin au masculin)
	F2078	Transformer un texte en appliquant des règles simples (féminin au masculin)
	F2079	Transformer un texte en appliquant des règles simples (féminin au masculin)
	F2080	Transformer un texte en appliquant des règles simples (féminin au masculin)
Comp15	M2025	Associer une unité usuelle à une grandeur (unité monétaire)
Niveau 1	M2026	Associer une unité usuelle à une grandeur (unité de poids)
	M2027	Associer une unité usuelle à une grandeur (unité de capacité)
	M2028	Associer une unité usuelle à une grandeur (unité de longueur)