

Environ un quart des élèves ont des performances indiquant une maîtrise satisfaisante des connaissances et compétences attendues en fin d'école primaire. À l'opposé, 15 % sont en difficulté : ils ne peuvent répondre qu'à des questions en lien avec leur expérience quotidienne ; 2 % environ sont en grande difficulté : bien que capables de répondre ponctuellement à quelques questions, ils ne maîtrisent aucune des connaissances et des compétences attendues en fin d'école. Entre les deux, 60 % des élèves se divisent en deux groupes égaux. Les moins performants répondent aux questions incluant des consignes simples qui leur permettent d'effectuer des correspondances terme à terme ; les autres atteignent un premier niveau de conceptualisation et peuvent exploiter des données organisées.

Les compétences des élèves en sciences expérimentales en fin d'école primaire

L'évaluation en sciences expérimentales en fin d'école, réalisée en juin 2007, avait pour finalité de mesurer les atteintes des objectifs fixés par les programmes de 2002 : « *L'enseignement des sciences et de la technologie à l'école vise la construction d'une représentation rationnelle de la matière et du vivant par l'observation, puis l'analyse raisonnée de phénomènes qui suscitent la curiosité des élèves.* »¹. Elle vise donc à donner non seulement des informations sur ce que savent les élèves (notion et lexique) mais aussi sur les compétences qu'ils ont développées (raisonnement, prélèvement...). Les programmes de 2002 rappelaient également que « *les compétences et les connaissances sont construites dans le cadre d'une méthode qui permet d'articuler questionnement sur le monde et démarche d'investigation.* » Aussi, en référence au programme, cette évaluation se compose-t-elle d'épreuves sous forme de questions à choix multiple (QCM) et de situations expérimentales destinées à approcher la démarche d'investigation développée par les élèves. De plus, des questionnaires de contexte, à destination des élèves, des enseignants et des directeurs d'école permettent, au-delà du simple constat d'une plus ou moins forte atteinte des objectifs et des attentes fixés par les textes officiels, de rapprocher les domaines cognitifs, non cognitifs et contextuels pour apprécier leur

mutuelle interaction et leurs effets sur les performances des élèves.

Évaluer des connaissances et des compétences : la démarche d'investigation

L'évaluation porte sur les champs de connaissances définis dans les programmes : « La matière », « Le ciel et la Terre », le « Monde construit par l'homme, l'Unité et diversité du vivant », « Le corps humain et l'éducation à la santé ». Pour simplifier le propos, les trois premiers champs seront regroupés sous le vocable « Physique » et les deux derniers sous le vocable « Sciences de la vie et de la Terre » (SVT). « L'éducation à l'environnement » étant transdisciplinaire, ce domaine n'est pas représenté dans cette évaluation. Les parties du programme qui « *désignent des champs du savoir pouvant, de façon optionnelle, servir de support à des activités d'investigation supplémentaires* »² n'ont pas été évaluées ; tel est le cas du champ « Énergie » et des manifestations de l'activité de la Terre (volcans, séismes).

Dans chacun des champs retenus, le questionnement couvre l'évaluation de trois compétences nécessaires aux acquisitions

1. 1 *Bulletin officiel* n° 1 du 14 février 2002, « Horaires et programmes d'enseignement de l'école primaire ».
2. Ibid.

en sciences expérimentales : « connaître », « prélever », « analyser ». En outre, des compétences développées dans le domaine des technologies de l'information et de la communication (TIC), parce qu'elles viennent « en complément de l'observation directe ou pour confronter les résultats aux savoirs établis »³, sont également évaluées. Mais, cette évaluation des TIC, de par le nombre restreint des situations retenues dans le protocole, ne peut en aucun cas être comparée au brevet informatique et internet.

Dans les programmes de 2002, il est indiqué que « l'enseignant sélectionne une situation de départ qui focalise la curiosité des élèves, déclenche leurs questions et leur permet d'exprimer leurs idées préalables. [...] Il amène à sélectionner les questions qui se prêtent à une démarche constructive d'investigation débouchant sur la construction des savoir-faire, des connaissances et des repères culturels prévus par les programmes. »

Pour évaluer la démarche d'investigation, trois situations ont été retenues. Elles ont été choisies de façon à pouvoir être reproductibles à l'identique dans le temps et étaient accompagnées d'un questionnaire permettant de mettre en évidence le comportement des élèves. Une seule situation était évaluée par école. Seuls quatre élèves passaient le module expérimental. Les contraintes techniques inhérentes à une évaluation sous forme papier ont permis de n'envisager que des aspects particuliers de la démarche d'investigation. Néanmoins, les trois modules expérimentaux étaient complémentaires.

Une échelle de performance

Pour formaliser les résultats, une échelle décrivant six niveaux de performance a été construite. Pour chaque groupe, elle précise le niveau global de maîtrise des connaissances et des compétences qu'elle permet également de décrire (voir l'encadré méthodologique, p. 6). La bande verticale gauche indique le pourcentage d'élèves de chaque groupe, la bande grisée horizontale l'étendue croissante des connaissances et des compétences maîtrisées du groupe 0 au groupe 5 : pour chaque groupe, la partie

bleue de la barre indique l'étendue des compétences supplémentaires qu'il maîtrise par rapport au groupe de niveau immédiatement inférieur.

Les niveaux de connaissances et de compétences des groupes d'élèves

Les élèves du **groupe 0** (2,4 % de la population), bien que capables de répondre très ponctuellement à quelques questions, ne maîtrisent aucune des connaissances et des compétences attendues en fin d'école primaire.

Les élèves du **groupe 1** (12,6 % de la population) réussissent, en physique comme en sciences de la vie et de la Terre (SVT), les items ayant un support visuel relevant d'une iconographie familière et relatifs à des situations de la vie quotidienne. En physique, leurs connaissances sont « pragmatiques ». Ils savent, par exemple, à quoi sert une boussole, quelle est la fonction du thermomètre ou encore que le vent et le soleil permettent au linge de sécher rapidement. En SVT, ils ont principalement des connaissances qui concernent le corps humain, plus particulièrement à propos de la circulation sanguine et des dangers représentés par l'usage du tabac.

Les élèves du **groupe 2** (30,9 % de la population) montrent un faible niveau d'abstraction en physique. Les items réussis, comme pour le groupe précédent, font référence à des situations de la vie familière. Les élèves commencent à montrer quelques connaissances, notamment à travers l'utilisation d'un lexique spécifique. En physique, ils connaissent les noms des saisons en fonction des dates ou le nom donné à la journée la plus longue de l'année. Ils ont acquis les principes élémentaires de sécurité électrique. Ils sont également capables d'analyser un schéma électrique simple et de différencier un circuit en série d'un circuit en dérivation. En SVT, la maîtrise du lexique concerne surtout le corps humain, en particulier la respiration à travers des termes tels que « inspiration » ou « expiration » exprimant l'aspect directement visible de la fonction. Dans le champ « Unité et diversité du monde vivant », le lexique connu concerne surtout les différentes phases du développement.

Les seules notions maîtrisées sont essentiellement en rapport avec le monde animal. Pour les élèves du **groupe 3** (31 % de la population), on note un très net élargissement des connaissances et des compétences, même s'il n'y a pas de maîtrise conjointe de la notion et du lexique s'y référant, notamment en physique. Les élèves atteignent un premier niveau de conceptualisation. Leurs connaissances touchent à tous les champs évalués. En physique, elles concernent les mouvements de la Terre et leurs effets pour l'homme. Dans le champ « La matière », ils sont capables de repérer que deux paramètres varient dans une « expérience » ou de différencier horizontalité et verticalité. En SVT, on note les premières connaissances maîtrisées ayant trait aux mouvements et aux muscles, à la digestion et aux modes de reproduction des animaux. Les élèves font la distinction entre végétal et animal. Ils savent exploiter des données organisées et ont acquis des compétences dans le domaine du prélèvement d'indices, qu'il soit simple ou complexe.

En physique, les élèves du **groupe 4** (13,1 % de la population) ont une compréhension fine d'énoncés décontextualisés. Ils ont des capacités d'anticipation et d'analyse. Le différentiel de connaissances avec les groupes précédents portent surtout sur le domaine « ombre et lumière » : les élèves savent que la lumière se propage de façon rectiligne et que l'ombre est une région de l'espace qui ne reçoit pas de lumière. En SVT, ils réussissent les items concernant le monde végétal sans que l'on puisse garantir pour autant qu'ils maîtrisent les notions correspondantes puisqu'ils ont le plus souvent recours à l'iconographie proposée dont ils savent extraire l'information utile. Ils sont également capables d'analyser une suite de chaînes alimentaires pour déterminer la nature d'un maillon précis.

En physique, les élèves du **groupe 5** (10 % de la population) sont capables d'effectuer des traitements complexes, en particulier, en mobilisant conjointement des compétences et des connaissances culturelles et scolaires. En SVT, ils ont une vision globale des caractéristiques du vivant dont les principaux fondements sont, d'une manière générale, maîtrisés.

3. Ibid.

Globalement, les résultats du lexique (66,7 %) sont supérieurs à ceux des notions (63 %). La compétence « Prélever » regroupe des situations dans lesquelles les élèves peuvent faire soit des prélèvements directs sans inférence (lecture de thermomètres, par exemple), soit des prélèvements multiples sur le document support. Les items relevant de la compétence « Analyser » requièrent pour leur part un traitement complexe de l'information au cours duquel l'élève est amené en particulier à faire des inférences (graphique 2).

... qui cachent des disparités selon le sexe...

Avec des taux de réussite supérieurs à 66 % en SVT, garçons et filles ont des résultats comparables. En physique, les garçons réussissent mieux que les filles : l'écart le plus important (4 points de pourcentage) est constaté dans le champ « Monde construit par l'homme » ; les champs « La matière » et « Le ciel et la Terre » présentent des écarts d'environ 3 points. En revanche, les filles ont, pour les « TIC », une réussite moyenne supérieure de près de 2 points.

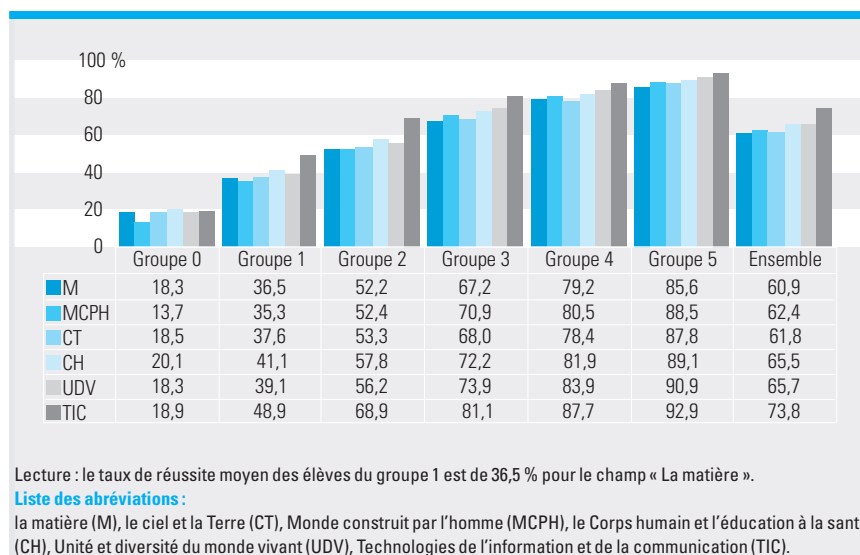
... le type d'école...

Globalement, les écarts « public/privé » varient de 2 à 4 points en faveur du privé suivant les champs. Il en va de même pour les écarts « public/éducation prioritaire (EP) » qui s'échelonnent de 6,3 à 9,4 points en faveur du public hors EP selon les champs, l'écart le plus important étant en « Unité et diversité du vivant ». Les courbes des résultats de chaque type d'écoles sont parallèles (graphique 3).

... l'âge

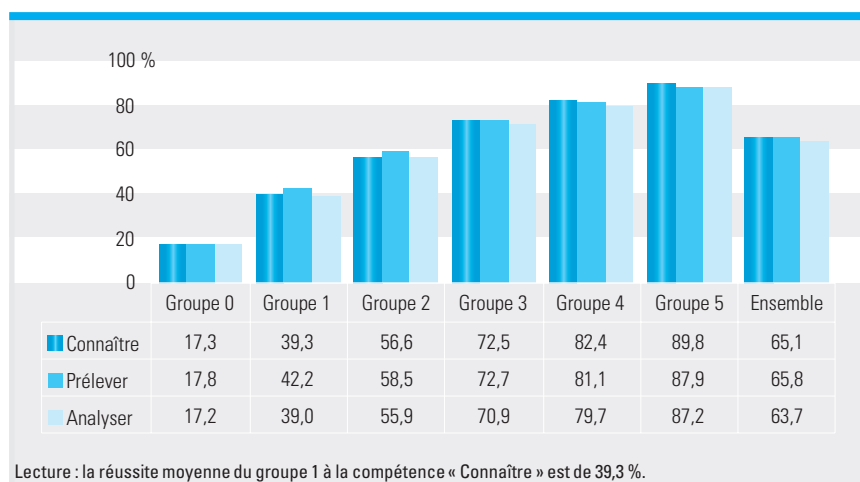
Quel que soit le champ, les écarts entre élèves « à l'heure » et « en retard » sont supérieurs à 14 points en physique et atteignent 16,8 points en SVT. En ce qui concerne les écarts entre élèves « à l'heure » et « en avance », le champ « Unité et diversité du vivant » est celui où les écarts sont les plus importants (11 points). C'est en « TIC » que les écarts sont les plus faibles (6 points). Dans les autres champs, les écarts varient de 8 à 10 points (graphique 4).

GRAPHIQUE 1 – Pourcentages de réussite par champ pour chaque groupe



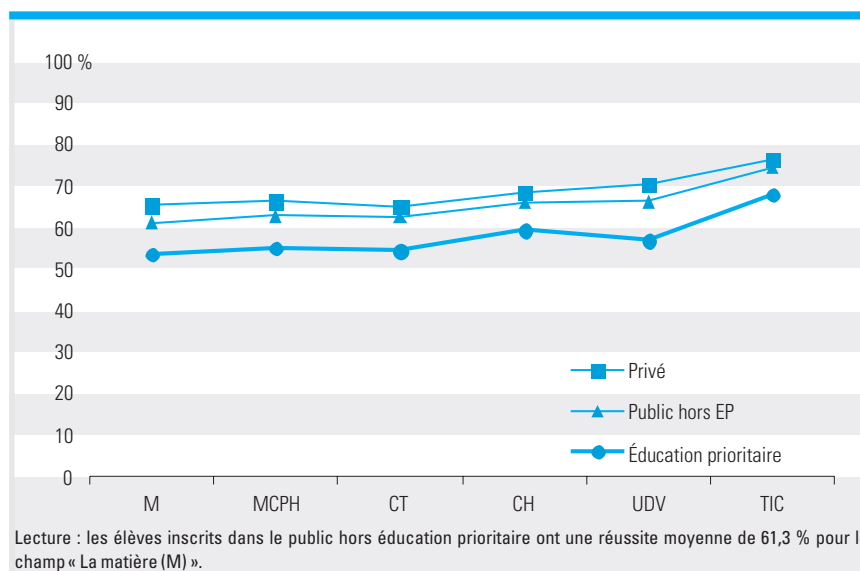
Source : MEN-DEPP

GRAPHIQUE 2 – Pourcentage de réussite par compétence



Source : MEN-DEPP

GRAPHIQUE 3 – Pourcentages de réussite par champ selon le type d'établissements



Source : MEN-DEPP

Une plus grande réussite en physique qu'en SVT

61 % des élèves maîtrisent le domaine de la matière, même si la notion d'évaporation demeure le concept le moins maîtrisé. Les faibles taux de réussite à certains items peuvent laisser penser que les élèves ont peu expérimenté, voire très peu utilisé, des instruments de mesure ; par exemple, pour mesurer une température.

Dans le domaine « Le ciel et la Terre », à la vue des résultats, on peut faire l'hypothèse que les enseignements de la mesure du temps ne sont pas liés aux phénomènes astronomiques concernés.

Dans le domaine « Le monde construit par l'homme », en électricité, les notions sont mieux maîtrisées que le lexique. Dès que les situations sont contextualisées, objet isolant ou objet conducteur introduit dans un circuit par exemple, les réussites (67 %) sont plus élevées que lorsqu'il s'agit de définir ce qu'est un isolant (55 %) ou un conducteur (39 %). Des manipulations sont sans doute faites mais les connaissances, tant sur le lexique que sur les notions, ne sont pas encore suffisamment mises en lien, ni entre elles ni avec les résultats des manipulations.

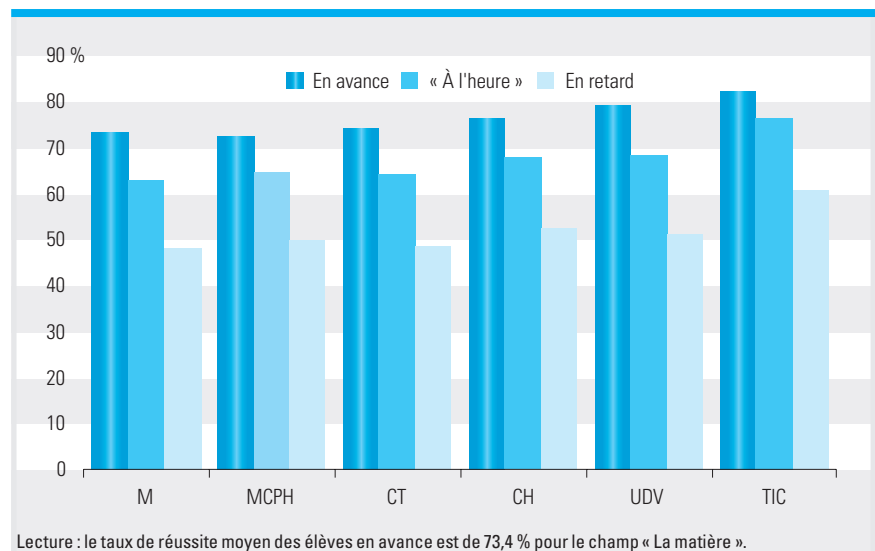
Dans le domaine « Le corps humain », lexique et notion ont des taux de réussite très variables. On note pour les items relatifs au corps humain, un taux de réussite de 66 %.

C'est à propos des fonctions de nutrition que l'on constate une grande amplitude des réussites. Les items relatifs à la digestion atteignent un taux moyen de réussite de 52,8 %, qui cache des disparités importantes. Au delà des principes anatomiques mieux connus, la connaissance des principes physiologiques apparaît lacunaire, puisque la fonction digestive n'est perçue que dans la dimension transit. La dimension transformation et absorption n'est pas du tout appréhendée par les élèves.

En revanche, les connaissances concernant la respiration sont assimilées par 68 % des élèves. Dans ce domaine, on constate également que le lexique (74 %) est mieux maîtrisé que l'aspect notionnel (65,5 %).

Pour la circulation sanguine, la maîtrise lexicale (74 %) se situe sensiblement au

GRAPHIQUE 4 – Pourcentages de réussite par champ selon l'âge



Source : MEN-DEPP

même niveau que pour la respiration (73 %). Dans ce domaine, les items relatifs aux aspects notionnels sont mieux réussis (80 %). Néanmoins, le rôle du sang est peu connu.

Dans le champ « Unité et diversité du vivant », s'agissant des caractéristiques du monde vivant, les réussites aux items qui concernent l'identification des organes et de leur rôle dans le monde végétal varient de 72,3 % à 78,8 %. Dès que les questions concernent une notion, on note une baisse significative des résultats (52,3 % à 65 %), ce qui confirme ce qui avait été observé à propos du corps humain.

Les notions de cycle de vie sont particulièrement bien perçues (plus de 80 % de réussite). Il n'en va pas de même pour les notions concernant la reproduction puisque l'on constate un taux global de réussite avoisinant les 50 %. Ces résultats méritent une attention particulière puisque, là encore, on observe un écart significatif entre les items relatifs au monde végétal (23 %) et ceux relatifs au monde animal (81,2 %).

Une structuration insuffisante des connaissances acquises à l'école

En physique, la fragilité du lien entre une notion et le lexique s'y référant montre la nécessité de renforcer la structuration des connaissances. Ceci est particulièrement souhaitable dans le domaine de la matière, en ancrant l'apport et l'utilisation du

lexique dans des situations permettant l'apprentissage des notions. Une simple « leçon de mots » ne peut suffire. À l'inverse, une suite de manipulations sans retour sur les notions sous-jacentes ne peut être efficace.

En SVT, le faible niveau de performances observé à propos du monde végétal suggère que, s'il existe des pratiques dans ce domaine, elles sont insuffisamment orientées vers une structuration des connaissances qui dépasserait la simple observation.

Le statut particulier des « TIC »

Les élèves maîtrisent les premières bases de la technologie informatique et ont une approche des principales fonctions d'un ordinateur (80 % de réussite moyenne) ; seule la fonction du scanner est moins connue (56 %). Les élèves disent adopter une attitude citoyenne face aux informations véhiculées (66 %), connaissent les fonctions correspondant aux icônes de la barre d'outils (89 %) et savent se documenter au moyen d'un produit multimédia (71 %).

La démarche d'investigation

Le premier module, qui abordait la matérialité de l'air, permettait de confronter les représentations initiales des élèves aux résultats d'une expérimentation qu'ils ne conçoivent pas, mais qui leur est proposée. Dans le cas de cette notion, qui se construit dans la durée,

la confrontation directe de l'élève à une situation expérimentale ne lui permet pas de la construire seul. Ce module montre que l'expérience permet des modifications (en positif comme en négatif) des prévisions, en grande partie guidées par les représentations des élèves, mais qu'elle ne peut se suffire à elle-même. L'expérience, sans traitement par l'enseignant, agit peu sur la compréhension de l'élève. La seule observation du phénomène est insuffisante et ne permet pas à l'élève de construire le concept. Dans le deuxième module, les élèves ne concevaient pas de protocole expérimental.

Ils devaient tâtonner à partir d'un dispositif, et interpréter des résultats pour comprendre un fonctionnement. 70 % d'entre eux ont fait des prévisions justes, 62 % ont réalisé des schémas corrects avec une légende et 14 % sans légende. Les élèves sont capables de choisir une anticipation correcte de ce qui va se passer sans pour autant être capables de la représenter. Après manipulation, ils sont capables à 87 % de choisir, parmi quatre propositions, la formulation qui décrit ce qu'ils viennent de vivre. Mais, lorsqu'ils doivent eux-mêmes produire (mettre en mots) une explication, l'écrit

reste un obstacle : ils ne sont plus que 42 % à y parvenir.

Dans le troisième module, les élèves concevaient un protocole expérimental, ce qui les amenait à prévoir, anticiper et concevoir une investigation, à partir de contraintes de matériel. Les trois quarts des élèves sont capables de concevoir un protocole expérimental. Néanmoins, on note une différence de réussite de dix points entre le dispositif prévu, représenté, et le dispositif effectivement réalisé.

**Gérard Brézillon et
Agnès Brun, DEPP B2**

Méthodologie

Le CEDRE

Le cycle des évaluations disciplinaires réalisées sur échantillon établit des bilans nationaux des acquis des élèves en fin d'école et en fin de collège. Il couvre les compétences des élèves dans la plupart des domaines disciplinaires en référence aux programmes. La présentation des résultats permet de situer les performances des élèves sur des échelles de niveau allant de la maîtrise pratiquement complète de ces compétences à une maîtrise bien moins assurée, voire très faible, de celles-ci. Renouvelées tous les six ans, ces évaluations permettront de répondre à la question de l'évolution du « niveau des élèves » au fil du temps.

Le calendrier des évaluations-bilans

2003 : compréhension écrite et orale (publiée)

2004 : langues étrangères (publiée)

2005 : attitudes à l'égard de la vie en société (publiée)

2006 : histoire-géographie (publiée)

2007 : sciences

2008 : mathématiques (en cours de publication)

2009 : reprise du cycle compréhension écrite et orale

L'échantillon

Un échantillon représentatif des écoles et des élèves inscrits en CM2 (4 156 élèves répartis dans 205 classes et 156 écoles) a été constitué au niveau national (établissements publics ou privés sous contrat de France métropolitaine). L'échantillon a été tiré dans la base statistique sur les établissements publics ou privés sous contrat de France métropolitaine (base de 1999-2000, complétée par celles de 2004-2005 ou 2005-2006 lorsque l'information est disponible, la base de 2005-2006 étant incomplète pour cause de grève administrative des directeurs d'écoles).

Format des questions

En raison de contraintes techniques fortes, la majorité des questions posées sont du format « questions à choix multiples » (QCM). Cette forme de questionnement a pour avantages d'éviter un travail de correction et d'automatiser la saisie des réponses des élèves, mais a pour inconvénient de ne pas permettre, en situation d'autonomie, l'évaluation des

compétences des élèves en production d'écrit. Les QCM ont été saisies de manière automatisée. L'ensemble de ces questions a été élaboré à partir des réponses que les élèves ont fournies lors d'une expérimentation réalisée en mai 2006.

Les cahiers tournants

Si chaque élève avait dû passer l'ensemble des situations proposées, deux heures d'évaluation par élève auraient été nécessaires. Pour limiter la passation à une heure pour chaque élève, les 165 situations d'évaluation des compétences ont été réparties en huit « blocs » agencés ensuite dans sept cahiers différents. Ce dispositif, permet d'estimer la probabilité de réussite de chaque élève à chaque item sans que chaque élève ait passé l'ensemble des items.

La construction de l'échelle de performances

L'échelle de performances a été élaborée en utilisant des modèles statistiques particuliers (modèles de réponse à l'item). Le score moyen correspondant à la performance moyenne des élèves de l'échantillon, a été fixé par construction à 250 et l'écart-type à 50. L'échelle n'a aucune valeur normative, la moyenne de 250 ne constitue en rien un seuil qui correspondrait à des compétences minimales à atteindre. Par analogie avec ce qui a déjà été fait pour d'autres évaluations-bilans, la partie la plus basse de l'échelle est constituée des scores obtenus par les 15 % d'élèves ayant les résultats les plus faibles. À l'opposé, la partie supérieure, constituée des scores les plus élevés, rassemble 10 % des élèves. Entre ces deux niveaux, l'échelle a été scindée en trois parties d'amplitude de scores égale correspondant à trois groupes intermédiaires.

Dans la théorie du modèle de réponse à l'item, les scores des élèves et la difficulté des items sont mesurés sur une même échelle, ce qui permet d'établir une correspondance entre les groupes d'élèves et les items répartis en ensembles de difficulté croissante.

On soulignera que les compétences évaluées en fin d'école et en fin de collège sont différentes, aucun élément commun ne permet de rapprocher les deux évaluations, il n'est donc pas légitime de comparer cette échelle avec celle du collège.

Pour en savoir plus

« Bilan histoire, géographie et éducation civique à la fin de l'école », *Note d'Information* 07.44, MEN-DEPP, décembre 2007.

F. Gibert et J.-M. Pastor, *Compréhension à l'écrit et à l'oral des élèves en fin d'école primaire – Évaluation 2003*, Les dossiers, n° 185, MEN-DEPP, juillet 2007.

« La maîtrise du langage et de la langue française en fin d'école primaire », *Note Évaluation* 04.10, MEN-DEP, octobre 2004.

www.education.gouv.fr

depp.documentation@education.gouv.fr

DIRECTION DE L'ÉVALUATION,
DE LA PROSPECTIVE ET DE LA PERFORMANCE
Directeur de la publication : Michel Quéré

Secrétaire de rédaction : Francine Le Neveu
Maquettiste : Frédéric Voiret
Impression : Ovation

DEPP, Département de la valorisation et de l'édition
61-65 rue Dutot – 75732 Paris Cedex 15
depp.diffusion@education.gouv.fr

Reproduction autorisée à la condition expresse de mentionner la source

ISSN 1286-9392