

note

évaluation

04.09
OCTOBRE

www.education.gouv.fr/stateval

Environ un quart des élèves semble avoir un bon ou un excellent niveau de maîtrise des compétences générales attendues en fin de collège : ils sont capables de mettre en œuvre des stratégies efficaces pour exploiter un document, de résumer, de synthétiser, d'argumenter, de mener un raisonnement complexe, d'utiliser la pensée hypothético-déductive. À l'opposé, 15 % s'avèrent n'avoir pratiquement aucune maîtrise de ces compétences ou une maîtrise réduite et sont en difficulté devant un texte complexe ou comprenant un vocabulaire peu courant ; 2% environ sont en grande difficulté : s'ils peuvent répondre ponctuellement à quelques questions, ils ne maîtrisent aucune des compétences générales évaluées.

Entre ces extrêmes, près de 60 % des élèves semblent avoir une maîtrise de ces compétences qui devrait les mettre à même de poursuivre une formation, même si une partie d'entre eux (de l'ordre du quart du total) est loin de posséder toutes les compétences générales attendues en fin de collège.

ministère
éducation
nationale
enseignement
supérieur
recherche



Les compétences générales des élèves en fin de collège

L'objectif des évaluations-bilans des acquis des élèves est de faire un point aussi objectif que possible sur les compétences et connaissances des élèves dans des domaines essentiels, à des moments-clés du cursus scolaire, et ceci en regard des objectifs fixés par les programmes.

Ces évaluations, destinées à rendre compte des résultats de notre système éducatif en établissant les différents niveaux de compétence atteints par les élèves au regard des objectifs fixés, seront renouvelées périodiquement. On disposera ainsi d'une appréciation de l'évolution des acquis des élèves que l'on s'efforcera de mettre en relation avec les caractéristiques du système et les politiques éducatives.

Les compétences générales des élèves en fin de collège

L'évaluation-bilan présentée ici a pour objectif d'évaluer des compétences générales ou procédurales (savoir-faire) attendues en fin de collège, compétences qui devront ensuite être réinvesties lors de la formation ultérieure, que celle-ci soit immédiate ou différée. Cette évaluation s'appuie sur les programmes du collège, mais n'est pas une évaluation directe de ce qui est défini dans ces programmes. En effet, si des documents officiels tels ceux du Conseil national des programmes¹ cherchent à « présenter le sens général des programmes et leurs principaux

objectifs en les structurant autour de trois grands pôles qui regroupent chacun plusieurs disciplines particulières : la maîtrise des langages, la culture des humanités, la culture scientifique et technique », il n'existe ni programme officiel ni définitions ni références établies concernant des compétences générales.

C'est pourquoi, on a repéré des compétences générales communes aux disciplines à partir d'une lecture croisée des programmes, mais on ne trouvera évidemment pas dans cette évaluation de réponses à des interrogations concernant des compétences relatives à une discipline particulière comme le français ou les mathématiques.

Les trois grands domaines de compétences

Trois grands domaines de compétences, gradués en fonction d'une complexité croissante des opérations mentales nécessaires pour les mettre en œuvre, ont été retenus :

- prélever l'information ;
 - organiser l'information prélevée ;
 - exploiter l'information de manière complexe.
- Ils sont ainsi définis et composés :

« **Prélever l'information** » est relatif à l'ensemble des moyens dont l'élève dispose pour accéder aux informations contenues dans les documents (textes, images, graphiques,

1. Conseil national des programmes, *Qu'apprend-on au collège*, CNDP et XO éditions, Paris, 2002.

cartes, plans...). L'information à prélever peut l'être de façon immédiate (information explicite), ou ne pas être immédiatement repérable et nécessiter le repérage de plusieurs éléments à combiner.

« **Organiser l'information prélevée** » consiste à trier, associer, choisir, reformuler, transposer d'un support à un autre, appliquer directement un principe, une règle, une méthode donnée à une nouvelle situation, effectuer des inférences simples à partir de vocabulaire inconnu, à partir d'indices textuels ou typographiques, d'accords éventuels.

« **Exploiter l'information de manière complexe** » est défini par le raisonnement déductif et la construction d'inférences complexes nécessaires à la compréhension du sens implicite d'un texte ou des relations de cause et de conséquence, au choix d'une argumentation, à la mise en œuvre d'une démarche expérimentale.

Pour chacun de ces grands domaines de compétences, on a cherché à définir une hiérarchie des différents niveaux de maîtrise, cette hiérarchie étant ensuite mise à l'épreuve, corrigée lors de l'expérimentation du protocole, puis confirmée lors de l'exploitation de l'enquête comme on le verra plus loin.

La conception des épreuves

Différentes composantes de ces trois domaines de compétences ont été envisagées dans quatre grands champs disciplinaires : français, mathématiques, histoire-géographie-éducation civique, sciences, en prenant en compte ce que ces composantes apportaient à la maîtrise de la compétence commune et non pas à la spécificité disciplinaire.

L'évaluation est ainsi organisée en situations réparties dans les quatre grands champs disciplinaires. Chaque situation est composée d'un document de travail, support de l'évaluation et de plusieurs questions. Chaque support d'évaluation permet de proposer de une à neuf questions, chacune d'elles ciblant un seul niveau de compétence. Au final, l'analyse a porté sur 55 situations regroupant 240 questions

ou items. La diversité des situations et des supports a été recherchée et les documents proposés sont d'une nature et d'une complexité adaptées aux élèves de fin de collège.

Une échelle de performances

Pour formaliser un constat de la maîtrise des compétences générales par les élèves en fin de collège, une échelle décrivant six niveaux de performance a été construite (voir le schéma p.3). Six groupes d'élèves correspondant à des niveaux différenciés de maîtrise des compétences ont pu ainsi être distingués. L'échelle de référence ainsi constituée permettra de voir si la répartition des élèves dans les différents niveaux de performance évolue dans le temps, lors des reprises successives de l'évaluation.

L'échelle précise la description du type et du niveau des compétences maîtrisées par les élèves de chaque groupe. La barre grisée symbolise l'étendue croissante de la maîtrise des compétences du groupe 0 au groupe 5.

Les niveaux de compétences des élèves

Une maîtrise différenciée des compétences générales

L'examen de l'échelle de performances et du pourcentage de réussite par compétence (graphique 1 p.4) montre la gradation des traitements à l'œuvre depuis la compétence la moins exigeante, le prélèvement de l'information, jusqu'à la plus exigeante, l'exploitation de l'information de manière complexe : ainsi, le prélèvement demande un repérage d'informations qui sont toutes données par le texte, alors que l'exploitation de l'information nécessite la mise en relation d'éléments fournis par le texte et d'informations qui lui sont extérieures. Or, les acquisitions que l'enseignement reçu au collège est censé apporter aux élèves supposent la

capacité de dépasser la simple compréhension de documents qui ne sont que les supports à partir desquels d'autres savoirs, d'autres connaissances et compétences, de niveau supérieur, seront développées. Pour acquérir ces nouveaux savoirs, il est indispensable que les élèves sachent organiser les informations contenues dans un document et, pour ce faire, il faut qu'ils sachent prélever les éléments essentiels contenus dans les documents qui leur sont présentés. Cet examen est ainsi très révélateur des acquisitions des différents groupes d'élèves et permet de fournir un indicateur précieux pour apprécier la situation des élèves de chacun de ces groupes au regard des compétences générales attendues en fin de collège.

Les élèves dont les performances les positionnent dans les groupes 0 et 1, réussissent rarement les items relevant des compétences les plus exigeantes. Ceux des groupes 2 et 3 peuvent réussir des items relevant de niveaux variés d'exigence. Un élève peut, par exemple, réussir une inférence à partir de la lecture du texte alors qu'il n'arrive pas à repérer les idées essentielles d'un texte pour le résumer. Les élèves des groupes 4 et 5 ont des performances élevées dans chaque compétence.

Comme on l'avait postulé, la hiérarchie des taux de réussite aux items relevant des différentes compétences est la même, pour chacun des groupes : les compétences de prélèvement d'information sont les mieux maîtrisées et celles d'exploitation de l'information les moins bien maîtrisées.

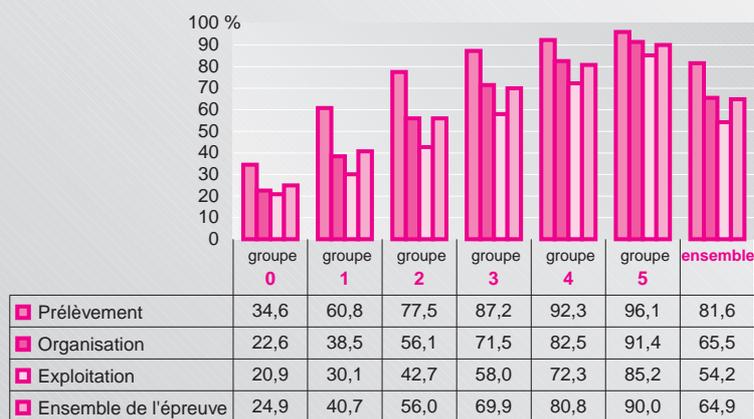
Six groupes aux compétences différenciées

Les élèves du **groupe 0** ne sont capables que très ponctuellement de répondre à une sollicitation. Leur meilleur taux de réussite est atteint aux items relatifs à la compétence « prélever l'information » (34,6 %). Pour les autres compétences, les taux de réussite aux items sont très faibles (autour de 20 %), quelles que soient les entrées prises en considération. On peut faire l'hypothèse que ces taux de réussite sont révélateurs de graves lacunes quant aux capacités les plus fondamentales de traitement de l'information écrite. On peut notamment s'interroger sur la possession par ces élèves des compétences de compréhension de lecture qui sont à la base de tous les traitements plus élaborés constitutifs des compétences plus complexes.

Avertissement

Étant donné que les compétences évaluées en fin d'école (*Note Évaluation 04.10*) et en fin de collège sont différentes, qu'aucun élément commun ne permet de rapprocher les deux évaluations, et que les jeunes non scolarisés en troisième (présentant sans doute un faible niveau de compétences) n'ont pas participé à l'évaluation fin de collège, il ne serait pas légitime de comparer cette échelle avec celle présentée dans la *Note Évaluation 04.10*.

Graphique 1 – Pourcentages de réussite par compétence pour chaque groupe



Les élèves du **groupe 1** ne montrent pas de difficultés aussi importantes que ceux du groupe 0. Avec 60,8 % de réussite moyenne aux items de « prélèvement de l'information », on peut les considérer comme des élèves fragiles du point de vue de cette compétence pour laquelle ils restent bien en deçà de la moyenne de réussite de l'ensemble des élèves (81,6 %). Cette fragilité est probablement à la source de leurs difficultés à faire preuve des compétences plus exigeantes. Ces élèves forment avec ceux du groupe 0 l'ensemble des élèves en grande difficulté que diverses études détectent tout au long de la scolarité au collège.

Les élèves du **groupe 2** représentent presque 30 % de la population. Leurs difficultés, loin d'être aussi aiguës que celles rencontrées par les élèves des deux premiers groupes, restent cependant importantes. Ces élèves maîtrisent la compétence « prélever l'information » (77,5 %) pour laquelle la marge de progression reste cependant importante. Bien que leur taux de réussite (56,1 %) pour les items de la compétence « organiser l'information prélevée » soit largement supérieur à celui du groupe 1, on ne peut considérer qu'ils atteignent un niveau de maîtrise suffisant de cette compétence. La compétence « exploiter l'information de manière complexe » est celle pour laquelle ils rencontrent les difficultés les plus sérieuses. Leur réussite est faible (42,7 %), n'atteignant les 50 % pour quasiment aucun aspect de cette compétence.

Les élèves du **groupe 3** rencontrent peu de difficultés relatives à la compétence « prélever l'information » dont ils réussissent les items à un niveau élevé (87,2 %). Ils maîtrisent assez bien la compétence « organiser l'information prélevée » et l'ensemble des éléments qui la composent (71,5 %). C'est la

compétence « exploiter l'information de manière complexe » qui leur pose problème. Ils ne la maîtrisent, ni globalement, (58 % de réussite aux items) ni dans la majorité de ses aspects. Toutefois, même s'ils sont encore peu assurés sur cette compétence, ils en réussissent essentiellement deux aspects : ils montrent une assez bonne compréhension des relations de cause et de conséquence (68,9 %) et parviennent relativement bien à tirer la conclusion d'une expérience (66,6 %).

La réussite des élèves des **groupes 4 et 5** est élevée, voire très élevée, pour tous les items des compétences évaluées. La maîtrise différenciée des compétences est encore respectée ici, mais les différences de performances relatives aux trois compétences s'amouindrissent, et ce davantage pour le groupe 5. La différence de performance entre les groupes 4 et 5 réside dans un écart moyen d'environ 9 points, écart qui s'explique par la différence de réussite aux items « d'organisation de l'information » et, surtout, aux items « d'exploitation de l'information ». D'une façon générale, les élèves du groupe 5 maîtrisent à un degré supérieur les trois compétences évaluées, notamment la compétence « prélever l'information ». C'est pour la compétence « exploiter l'information de manière complexe » que le groupe 5 voit son niveau de performance le distinguer le plus du groupe 4 (écart de 12,9 points).

La réussite supérieure des élèves du groupe 5 relativement à ceux du groupe 4 est particulièrement éclatante pour les items qui évaluent les formes de raisonnement complexe telles celles liées à la mise en œuvre d'une démarche expérimentale et à la capacité d'accepter ou de rejeter une procédure permettant la résolution de problème. Cette dernière entrée, mettant en jeu

le raisonnement hypothético-déductif, reste cependant la plus difficile à acquérir et à maîtriser, même pour les élèves les plus performants à l'issue du collège.

Des performances qui varient

Selon l'académie...

Un score moyen attendu de chaque académie est calculé en tenant compte de sa structure sociale ². Ainsi, l'académie de Clermont-Ferrand a un score attendu de 250 et un score observé de 255. Le score observé est le score moyen obtenu par les élèves de l'échantillon pour cette académie à l'évaluation. Les académies situées au-dessus de la droite ont un score observé supérieur à leur score attendu. Les académies situées en dessous de la droite ont un score observé inférieur à leur score attendu. Les académies situées sur la droite ont un score observé égal à leur score attendu (*graphique 2 p.5*).

... les conditions de scolarité...

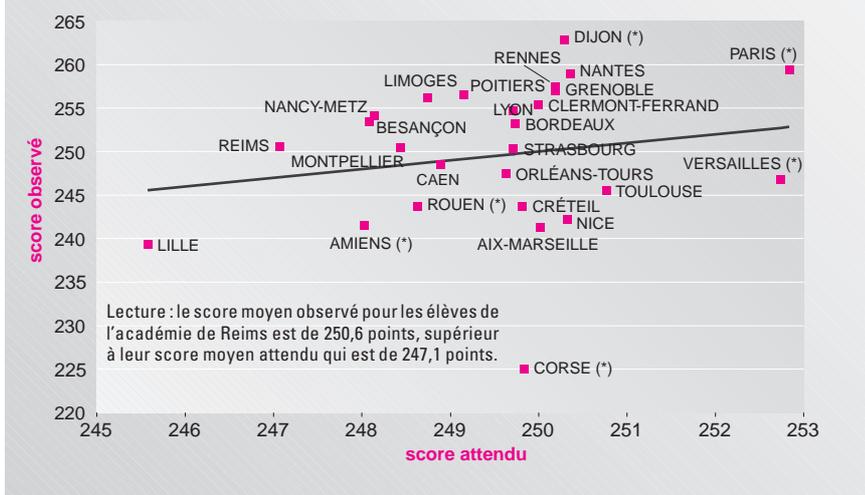
L'analyse des résultats en fonction des types de zone laisse apparaître des constats du même type que ceux habituellement observés dans les évaluations menées par la Direction de l'évaluation et de la prospective, à savoir des niveaux de performance sensiblement inférieurs pour les élèves situés en ZEP/REP comparés à ceux observés pour l'ensemble des élèves. En ZEP/REP, 25,7 % des élèves font partie des groupes 0 et 1 (soit presque 2 fois plus que dans l'ensemble de la population) alors qu'ils ne sont que 15,2 % dans les groupes 4 et 5 (soit 1,7 fois moins que dans l'ensemble de la population et 2 fois moins pour le seul groupe 5).

...et le cursus scolaire

Les performances des élèves sont très différenciées selon le cursus scolaire suivi à l'école et au collège. D'une façon générale, pour les catégories d'élèves qui ont subi un

². Cependant, la variable PCS a été très mal renseignée. Pour certaines académies, le taux de non-réponses peut atteindre les 80 %. Pour calculer les scores attendus, on a donc tenu compte à la fois de la structure sociale de l'académie dans l'échantillon et de la structure sociale réelle. Pour la plupart des académies ces deux structures sont équivalentes sauf pour six d'entre elles : Paris, Versailles, Dijon, Rouen, Amiens et la Corse. Pour ces académies repérées par un astérisque (*graphique 2*), les scores attendus sont donc à considérer avec prudence.

Graphique 2 – Situation des académies au regard des performances attendues



redoublement, la proportion d'élèves de chacun des trois groupes supérieurs est inférieure à celle de l'ensemble de la population tandis qu'elle est supérieure pour chacun des trois groupes les plus faibles. Cette tendance est d'autant plus marquée que le redoublement a été précoce. Ainsi, les élèves qui ont redoublé sont nettement moins nombreux à atteindre les groupes 4 et 5, et cela d'autant plus que le redoublement a eu lieu précocement. Ils sont *a contrario* beaucoup plus nombreux à se situer dans les groupes 0 et 1, là aussi d'autant plus que le redoublement a été précoce. En effet, les élèves qui ont redoublé à l'école primaire sont cinq fois plus nombreux à être en difficulté et en grande difficulté (41,6 % dans les groupes 0 et 1) que ceux qui n'ont jamais redoublé, alors que lorsqu'ils ont redoublé la troisième, ils ne sont que 10,7 % dans cette situation. En revanche, on observe l'effet in-

verse en ce qui concerne les groupes 4 et 5 : les élèves qui ont redoublé en primaire sont douze fois moins nombreux dans ces groupes (3 % contre 35,1 %) alors que ceux qui ont redoublé la troisième le sont seulement 2,5 fois (14,6 % contre 35,1 %).

Que nous apprennent ces résultats par rapport aux attentes du collège ?

Rappelons que l'ensemble de cette évaluation-bilan correspond à un premier état des niveaux de compétences des élèves en fin de collège. Un de ses objectifs est de permettre un suivi de l'évolution de ces niveaux de compétences au cours du temps. Les élèves des groupes 4 et 5 (25,9 % de la population) ont des performances qui peuvent permettre de considérer qu'ils maîtrisent de

façon très satisfaisante ou satisfaisante l'ensemble des compétences générales attendues en fin de collège. Les groupes 0 et 1 montrent des élèves en difficulté (groupe 1) et en grande difficulté (groupe 0). Leur proportion peut être évaluée à 15 % de la population de troisième. Les élèves du groupe 2 (29,6 %) sont des élèves qui, s'ils maîtrisent la compétence de prélèvement, n'atteignent que partiellement les compétences d'un niveau d'exigence plus élevée retenues par l'évaluation-bilan. Les élèves du groupe 3 (29,5 %) maîtrisent les compétences de premier et deuxième niveau d'exigence, « prélever l'information » et « organiser l'information prélevée », mais butent sur celle du niveau le plus élevé, « exploiter l'information de manière complexe ».

Le groupe 2 mérite une attention particulière. Au regard de leurs performances, l'hypothèse selon laquelle les élèves de ce groupe ne possèdent pas les compétences attendues en fin de collège, et n'ont donc pas atteint les objectifs des programmes, est vraisemblable. Ces premiers constats doivent permettre de s'interroger sur la question du seuil à partir duquel les objectifs du programme sont atteints : l'échelle construite propose un échelonnement des réussites des élèves, mais elle ne fournit pas une norme au-delà de laquelle les exigences du programme seraient atteintes et en deçà de laquelle, elles ne le seraient pas. Elle illustre surtout l'hétérogénéité des élèves dans la maîtrise des compétences générales attendues en fin de collège par les programmes.

C'est pourquoi, il faut sans doute s'interroger sur deux autres questions qui sont liées : les objectifs des programmes constituent-ils un optimum, un idéal à atteindre par les meilleurs, un objectif d'excellence que se propose l'École, et si tel est le cas, quel « socle de compétences » doit être considéré comme nécessaire en fin de collège (c'est-à-dire en fin de scolarité obligatoire), à la fois pour poursuivre immédiatement ou plus tard une formation et pour s'insérer harmonieusement dans la vie professionnelle et sociale.

Par ailleurs, cette évaluation-bilan vient confirmer ce qui était déjà perceptible lors des évaluations internationales³ à savoir que nos élèves semblent plus à l'aise dans le prélèvement d'information (compétence considérée comme de bas niveau) alors

Graphique 3 – Répartition de la population sur l'échelle de compétences générales en fonction des cursus élémentaire et secondaire



Lecture : la population qui a redoublé la sixième se répartit sur l'échelle de compétences générales comme suit : - 3,6 % groupe 0 - 27,1 % groupe 1 - 42,7 % groupe 2 - 22,8 % groupe 3 - 3,1 % groupe 4 0,7 % groupe 5.

3. Voir « Les élèves de 15 ans – Premiers résultats d'une évaluation internationale des acquis des élèves (PISA) », Note d'Information 01.52, MEN-Direction de la programmation et du développement, décembre 2001.

qu'ils éprouvent des difficultés dans l'exploitation complexe de l'information contenue dans un document (compétence considérée comme de haut niveau).

Ce constat, comme celui de la grande hétérogénéité des compétences des élèves sortant du collège, implique :

– en amont, une réflexion sur la prise en compte des difficultés de ces élèves dès l'entrée en sixième, puisque l'on sait qu'elles sont

déjà constituées à ce moment et ne feront que se confirmer dans la mesure où la construction des compétences de plus haut niveau s'étaye sur celles du niveau le plus bas ;
– en aval, une réflexion sur la prise en compte des difficultés de ces élèves à l'entrée en seconde, qu'elle soit professionnelle ou générale et technologique.

Dans ces deux domaines les outils d'évaluation diagnostique mis à la disposition des en-

seignants, à travers la « banquutils »⁴, devraient permettre une meilleure détection des difficultés rencontrées par les élèves et donc une meilleure prise en compte de ceux-ci.

**Laurence Dauphin et
Bruno Trosseille, DEP C1**

4. <http://www.banquutils.education.gouv.fr>

Méthodologie

L'échantillon

Pour répondre aux finalités du dispositif, deux échantillons représentatifs des collèges et des élèves inscrits en troisième ont été constitués aux niveaux national et académique.

Pour une représentativité académique, trente élèves ont été tirés aléatoirement dans dix-sept établissements par académie. Pour une représentativité des élèves et des classes au niveau national, et pour travailler sur les « effets-classes », on a constitué un deuxième échantillon constitué de classes entières au sein d'un même établissement. Au total, le rassemblement de ces deux échantillons constitue un ensemble de 23 500 élèves, 552 collèges et 411 classes entières.

Les échantillons ont été tirés dans la base centrale des établissements publics ou privés sous contrat de France métropolitaine (base de 2002-2003). En raison notamment des mouvements sociaux du printemps 2003, seules les réponses de 18 222 élèves ont pu être analysées, ce qui a nécessité un redressement de l'échantillon.

On observera qu'en fin de collège la population ne comprend pas, contrairement à celle de l'école primaire, la totalité d'une promotion car certains élèves, généralement les plus âgés de ceux qui ont des difficultés, sont déjà sortis du système scolaire et ne se retrouvent donc pas dans nos échantillons.

Contraintes de la situation d'évaluation

En raison de contraintes techniques fortes (la taille de l'échantillon de départ interdisant une possibilité de correction par des experts), toutes les questions 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 l'évaluation des compétences des élèves en production d'écrit, en situation d'autonomie.

Par ailleurs, si chaque élève avait dû passer l'ensemble des situations proposées, huit heures d'évaluation par élève auraient été nécessaires. Pour limiter la passation à deux heures pour chaque élève, les cinquante-cinq situations ont été constituées en treize « blocs » répartis ensuite dans seize cahiers différents selon un plan expérimental, organisant un tuilage des blocs. Ce dispositif, couramment utilisé dans les évaluations-bilans, notamment les évaluations internationales, permet d'estimer la probabilité de réussite de chaque élève à chaque item sans que chaque élève passe 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 de compétences générales, correspondant à la performance moyenne des élèves de l'échantillon, a été fixé par construction à 250 et l'écart type à 50. Cela implique qu'environ deux tiers des élèves ont un score compris entre 200 et 300. Mais cette échelle, comme celle de l'enquête PISA, n'a aucune valeur normative et, en particulier, la moyenne de 250 ne constitue en rien un seuil qui correspondrait à des compétences minimales à atteindre.

Sur la base de constats fréquemment établis dans les différentes évaluations antérieures de la DEP, qui montrent que 15 % des élèves peuvent être considérés comme en difficulté en fin de scolarité obligatoire, 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 240 items répartis en ensembles de difficultés croissantes. Les compétences de chacun des groupes d'élèves ont pu ainsi être caractérisées, puis analysées.