

Préambule : comment évaluer les acquis des élèves ?

Nicole Braxmeyer

Professeure chargée d'études au bureau de l'évaluation des politiques éducatives et des expérimentations
Direction de l'évaluation et de la prospective

Les évaluations-bilans des acquis des élèves conduites par le bureau des évaluations et des outils pour le pilotage pédagogiques (DEPP-B2) ont pour objectif de faire le point sur les connaissances et les compétences des élèves dans des disciplines (domaine cognitif) ou des attitudes (domaine conatif), à des moments-clés du cursus scolaire, en regard des objectifs fixés par les programmes d'enseignement. Ces évaluations n'ont pas valeur de délivrance de diplômes, ni d'examen de passage ou d'attestation de niveau. Elles donnent une photographie à un instant « t » de ce que savent et savent faire les élèves à la fin d'un cursus scolaire, en ce sens il s'agit bien d'un bilan. Elles sont passées auprès d'un échantillon statistiquement représentatif de la population scolaire concernée. Le dispositif n'est pas conçu pour donner des informations sur un élève particulier, une classe ou un établissement scolaire. Destinées à être renouvelées périodiquement, ces évaluations-bilans permettent de disposer d'un suivi de l'évolution des acquis des élèves dans le temps. Elles apportent aussi l'éclairage qui intéresse tous les niveaux du système éducatif des décideurs aux enseignants sur le terrain, en passant par les formateurs : elles informent, bien sûr, sur ce que savent et savent faire les élèves à la fin d'un cursus ; elles éclairent sur l'attitude et la représentation des élèves à l'égard de la discipline ; elles interrogent les pratiques d'enseignement au regard des programmes et elles contribuent à enrichir la réflexion générale sur l'efficacité et la performance de notre système éducatif.

La réalisation de ces évaluations, proposées aux élèves, quels que soient les cheminements pédagogiques rencontrés au cours de leur apprentissage et avec des enseignants différents, nécessite plusieurs étapes.

Pour ce qui est de la présente évaluation, les épreuves ont été construites à partir d'un cadre de référence établi selon les exigences des programmes d'enseignement en vigueur en 2004 : du cycle III pour le premier degré et de la sixième à la troisième pour le collège. Ces exigences se déclinent en termes de connaissances que les élèves doivent acquérir et de compétences qu'ils doivent maîtriser à la fin d'un cursus scolaire. Les situations d'évaluations et les supports proposés aux élèves sont de difficultés variables, objets de questions indépendantes les unes des autres afin qu'une question ne constitue pas un frein ou, a contrario, un élément qui induise la réponse à la question qui la suit. Le matériel documentaire qui accompagne ces

évaluations est constitué de supports propres aux disciplines concernées : textes, cartes, documents iconographiques, graphiques, croquis, schéma, etc., en usage dans l'enseignement.

Les questions posées aux élèves sont du format questions fermées et se déclinent en « questions à choix multiples » (QCM) ou en vrai/faux. En effet, pour ces évaluations à grande échelle, la taille de l'échantillon représente une double contrainte : celle du nombre d'items proposés aux élèves et celle du traitement des réponses. Le format de questions fermées autorise la saisie automatique des réponses des élèves (réponses cochées, chiffres reportés ou entourés) car la taille de l'échantillon et le nombre d'items ne permettent pas d'envisager les corrections par des experts. Cependant, pour le collège (voir infra), des questions dites ouvertes ou construites ont été aussi proposées pour évaluer des connaissances et des compétences à l'occasion de productions écrites courtes ou élaborées.

À la fin du processus de conception, c'est près de dix heures d'évaluation qui auraient été nécessaires pour que les élèves répondent à la totalité des items produits. Pour limiter le temps de la passation pour chaque élève tout en proposant la totalité des items construits à l'échantillon désigné, un dispositif spécifique, utilisé couramment dans ce type d'évaluations, a été utilisé. Les items ont été répartis dans des blocs (16 pour l'école comme pour le collège) ; ces blocs ont été distribués selon un plan préétabli dans des cahiers (20 pour l'école, comme pour le collège) d'évaluations différents (« cahiers tournants »), chaque cahier proposant à chaque élève deux fois 50 minutes d'évaluation disciplinaire (plus le questionnaire de contexte). De ce fait, les élèves d'une classe n'avaient pas tous des cahiers identiques.

La saisie des réponses des élèves est automatisée, ce qui n'est guère d'un usage habituel dans le domaine scolaire. Pour le collège, la correction des questions expertes (questions ouvertes) par des professeurs s'est effectuée par télécorrection via une application informatique spécifique élaborée par la DEPP.

Sur le plan méthodologique, le bureau de l'évaluation, des politiques éducatives et des expérimentations s'appuie sur les concepts et les méthodes de la mesure en éducation et de la psychométrie reconnus et utilisés internationalement dont l'objet est de construire une mesure fiable, c'est-à-dire valide et fidèle¹. Pour ce qui concerne l'aspect disciplinaire, les situations d'évaluation sont basées sur les préconisations des programmes d'enseignement lesquels s'inspirent de courants didactiques et pédagogiques reconnus.

1. Laveault et Grégoire, 1997, Linn, 1989, Dicks et al. 1994

Les acquis des élèves en fin de collège

Michel Braxmeyer

Professeur agrégé de géographie, chargé d'études

Nicole Braxmeyer

PLP lettres-histoire, chargée d'études

Séverine Dos Santos

Chargée d'études statistiques,

Bureau des évaluations et des outils pour le pilotage pédagogique

Direction de l'évaluation, de la prospective et de la performance

Les outils d'évaluation des acquis des élèves sont réalisés à partir d'un cadre de référence conforme aux contenus des programmes et aux évaluations-bilans conduites par la DEPP. L'analyse statistique des réponses fondée sur le modèle de réponse à l'item (MRI) donne une première photographie de ce que savent et savent faire les élèves en fin de collège. Elle permet de dégager des profils d'élèves indiquant des niveaux d'acquisition de connaissances et de compétences. Les résultats révèlent une situation positive dans la lecture documentaire et dans l'appropriation des connaissances listées pour le diplôme national du brevet. Ils peuvent alors être confrontés aux attendus de fin de collège et interroger sur les pratiques pédagogiques en cours pour les atteindre. Ils interrogent sur le statut et l'usage du matériel documentaire dans les apprentissages, sur le passage à l'écrit dans l'évaluation de la discipline et sur l'apport de l'évaluation-bilan au regard des pratiques habituelles. L'opinion des élèves sur leur relation à la discipline est un bon indicateur de l'image attractive qu'elle dégage.

Comme annoncé dans le préambule, la conception des épreuves s'appuie sur les compétences, le matériel documentaire et les contenus des programmes d'enseignement. Une première analyse des résultats permet de dégager des profils d'élèves en termes d'acquisition de connaissances et de compétences qui permettra, par la suite, des comparaisons dans le temps. Ces résultats doivent être mis en regard avec les attendus de fin de collège (objectifs fixés par l'institution), certaines pratiques enseignantes et l'image que se font les élèves de la discipline.

LA CONCEPTION DES ÉPREUVES DE L'ÉVALUATION

Une première année a été consacrée à la réflexion préalable (contenu, choix...) et à la conception d'une version expérimentale de l'épreuve. Ce travail a été conduit par une équipe d'enseignants concepteurs (14 enseignants venant de différentes académies de métropole) qui a pu mettre au service du projet non seulement sa capacité et sa compétence à s'interroger sur l'évaluation mais aussi sa connaissance des pratiques sur le

terrain. Un comité scientifique veillait à la rigueur des travaux. Une première épreuve d'expérimentation a eu lieu à la fin de cette première année. Après une phase de correction et de réajustement, l'évaluation a été passée en mai 2006 auprès d'un échantillon représentatif de 197 collèges et 5 856 élèves. L'ensemble de l'épreuve comportait 1 035 prises d'information regroupées par la suite en 370 items (318 QCM et 58 questions ouvertes). Le principe des « cahiers tournants » a permis de proposer une grande variété de situations d'évaluation, une variété de tâches à effectuer et l'éventail du matériel documentaire utilisé couramment dans la discipline.

Les compétences visées

Les compétences permettant de cerner les acquis des élèves ont été retenues selon les finalités assignées à l'enseignement de l'histoire, de la géographie et de l'éducation civique préconisées en introduction des programmes de collège de la sixième à la troisième :

- des finalités intellectuelles : ces disciplines contribuent au collège « à former l'intelligence active » par l'apprentissage puis la maîtrise des « opérations élémentaires de la construction du

savoir » telles que : lire et identifier, reconnaître et nommer, puis organiser ce que l'on a appris à reconnaître (classer et mettre en relation) et enfin donner du sens ;

– des finalités civiques : « comprendre le monde et agir en personne responsable [...] exigent la connaissance du monde dans sa diversité et son évolution » ;

– des finalités culturelles : « donner aux élèves une vision du monde et une mémoire » (introduction des programmes d'histoire, géographie et éducation civique au collège).

La terminologie utilisée fait donc références aux définitions en vigueur avant le décret de juin 2006 sur le socle commun de connaissances et de compétences.

En histoire-géographie et éducation civique à partir des contenus des programmes et du corpus de connaissance, trois grandes compétences ont été dégagées :

1. Identifier (lire et identifier, reconnaître et nommer) décliné en trois composantes :

– localiser, énumérer, contextualiser : situer dans le temps et dans l'espace, lister les éléments d'un espace, d'une période, replacer un événement, un fait dans son contexte, etc. À travers cette composante, entre autres, est évalué ce que l'élève a mémorisé ;

– décrire : nommer par le vocabulaire simple ou spécialisé et par le vocabulaire notionnel ;

– sélectionner une ou des informations implicites ou explicites dans des documents ou dans sa mémoire.

2. Traiter l'information (organiser) relative à une situation, décliné en deux composantes :

– classer, hiérarchiser ;

– mettre en relation, comparer des informations, des données en croisant

différents langages ; établir des liens dont causes-conséquences ;

3. Interpréter une situation (donner du sens), décliné en trois composantes :

– généraliser : dégager une idée, donner un titre, donner un nom à un critère, résumer, proposer une synthèse...

– argumenter : expliquer un événement, un fait ; justifier une affirmation ; illustrer avec des exemples ; distinguer une information d'une opinion ; critiquer une information, un document...

– réaliser : utiliser le matériel documentaire et ses techniques, reformuler des informations en utilisant différents langages (rédaction, construction de cartes, de schémas, de croquis, de légende...).

Ces trois grandes compétences, bien entendu, ne sont pas hermétiquement étanches entre elles, ni hiérarchisées. Leur présentation distincte tente de clarifier des opérations intellectuelles demandées à l'élève dans l'objectif d'un apprentissage global. Dans le cadre d'une évaluation-bilan, cette présentation permet de positionner l'élève selon différents niveaux d'acquisition dans ces trois compétences.

Le matériel documentaire utilisé

L'enseignement de l'histoire, de la géographie et de l'éducation civique s'appuie sur l'initiation, la pratique, puis la maîtrise par l'élève des différents langages que représentent les outils de la discipline : la lecture et l'observation de textes écrits – authentiques ou informatifs voire de fiction – d'images (documents patrimoniaux, documents iconographiques reproduits), de cartes et de graphiques. Le document, étudié en tant que tel est objet d'apprentissage mais il est aussi support pour d'autres

apprentissages. (« L'enseignement de l'histoire et de la géographie doit initier à la pratique de différents langages », programme de sixième). Les situations d'évaluation s'appuyaient donc sur l'observation et la lecture de documents divers de complexité variée, pour mettre en œuvre des opérations de recherche d'information, de confrontation et d'interprétation.

Les situations d'évaluation élaborées à partir de thèmes fédérateurs

Il était impossible de prétendre à l'exhaustivité et de balayer la totalité du programme de collège. Cette compilation aurait-elle d'ailleurs un sens, voire un intérêt ? (Les enseignants étant appelés dans les documents d'accompagnement des programmes à choisir : « il est impossible de tout dire : enseigner c'est choisir » programme de sixième). Des choix ont été opérés en retenant des thématiques disciplinaires qui traversent les programmes de collège à partir desquelles les situations d'évaluation ont été créées : sept thèmes en histoire (religion, civilisation/société, expansion/conquêtes, organisation politique/construction de la République, guerre, révolution, démocratie/totalitarisme) ; cinq thèmes en géographie (répartition, paysages, échanges et flux, territoire, contraste de développement) ; quatre thèmes en éducation civique (citoyenneté/démocratie, liberté/droit/justice, solidarité/égalité, droits/devoirs de la personne). Un thème « repères » proposait des situations pour chacune des trois disciplines.

UNE PREMIÈRE PHOTOGRAPHIE DES RÉSULTATS

Cette évaluation ne permet pas de fournir des résultats à des fins de comparaisons telles que « *les élèves savent plus de choses en histoire qu'en géographie ou en éducation civique* ». Il en est de même pour les différents espaces, périodes ou thèmes questionnés. Cela supposerait des épreuves construites terme à terme à cet effet : équivalence sur les exigences des compétences et composantes, nombre d'items comparables, nombre de questions comparables, types de support comparables (en nombre et en niveau de difficultés). La construction des évaluations a été faite sur le principe que « *l'histoire et la géographie n'ont pas seulement des démarches intellectuelles communes, les contenus des programmes permettent de multiples rencontres qu'il est indispensable de*

favoriser » (introduction au programme de cinquième).

Les résultats livrés ici donnent une première photographie des acquis des élèves en fin de collège. L'analyse fine des résultats fera l'objet d'une publication ultérieure à paraître en 2008.

Le pourcentage moyen de réussite des élèves à l'ensemble de l'évaluation est de 58,6 %. C'est lorsqu'il faut « identifier, prélever l'information » que la réussite est la plus élevée (61,5 %). Cette compétence suppose de travailler sur les différents types de documents proposés. À l'intérieur de cette compétence, le pourcentage de réussite obtenu aux questions portant plus directement sur les connaissances (« localiser et contextualiser ») est le plus fort de l'ensemble de l'évaluation (62,2 %).

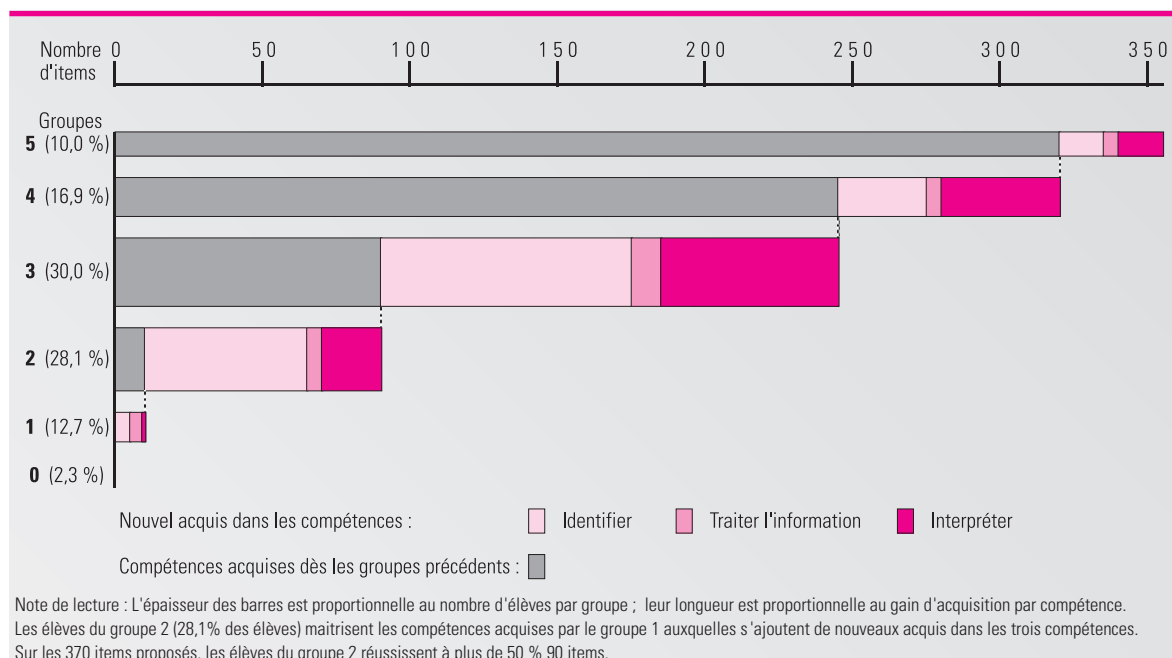
Lorsqu'il faut « traiter l'information » ou « interpréter », les pourcentages de réussite moyens obtenus par l'ensemble des élèves sont moins élevés (54,7 % et 55,2 %).

Dans la compétence « interpréter » lorsqu'on demande aux élèves de « produire, réaliser », le pourcentage moyen de réussite aux items qui l'évaluent n'est que de 46,9 %. Ce taux est le plus faible. Il faut préciser qu'il s'agissait de passer à la production écrite sous forme essentiellement de questions ouvertes.

Le modèle statistique utilisé par la DEPP (et par les évaluations internationales auxquelles participe la France telles que PISA et PIRLS) pour traiter les résultats permet de dégager six groupes gradués d'élèves. Pour chacun des groupes, on peut dresser un profil des acquis en termes de savoir et de savoir-faire, sachant que chaque groupe maîtrise mieux les compétences acquises par le groupe précédent auxquelles s'ajoute l'acquisition de nouvelles compétences (et ainsi de suite) (*graphique 1*).

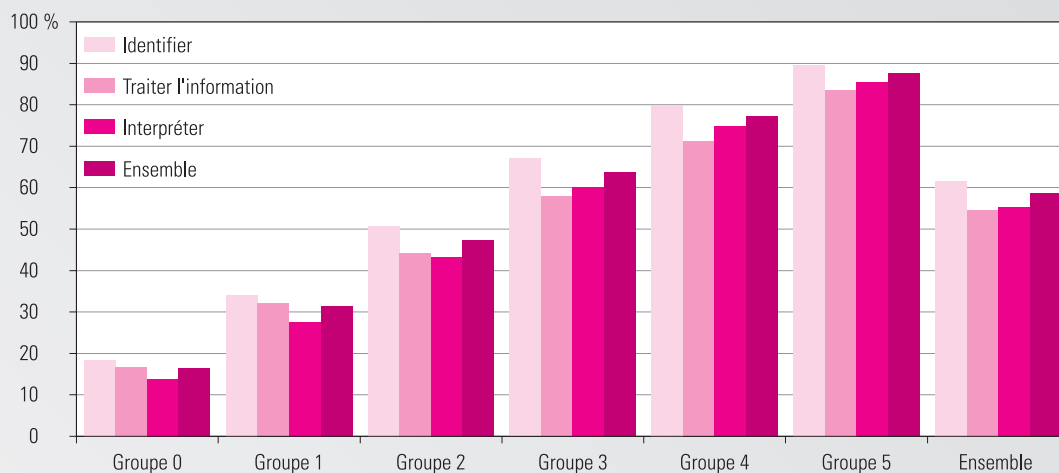
Cette manière de présenter les résultats met en évidence ce que

Graphique 1 – Échelle de progression dans les compétences par groupe d'élèves



Source : MEN-DEPP

Graphique 2 – Pourcentages de réussite par compétence pour chaque groupe



Lecture : Le groupe 3 obtient une réussite de 63,7% à l'ensemble des items et 60,2% à ceux de la compétence Interpréter.

Source : MEN-DEPP

savent faire les élèves, en terme de niveau d'acquisition, par groupe. Elle permettra des comparaisons dans le temps sur l'évolution de l'acquis de ces compétences. Elle ne saurait masquer les manques. Cette approche est plus éclairante que ne peuvent l'être les seuls pourcentages moyens de réussite, qu'ils soient globaux ou items par items (comment interpréter un pourcentage moyen de réussite ? par rapport à quelle référence ?).

Un petit groupe d'élèves (**2,3 % = groupe 0**) n'est capable que de réponses ponctuelles avec pourcentage moyen de réussite de 16,5 % seulement (*graphique 2*). Ces élèves ont très peu de connaissances concernant la discipline (notamment espace et temps), mais ils ont aussi des lacunes en compréhension

de l'écrit (quel que soit le support) qui empêchent le traitement et l'interprétation de l'information écrite.

12,7 % des élèves (groupe 1) ont un pourcentage de réussite à l'ensemble de l'évaluation de 31,5 %. Ces taux s'élèvent à 34 % pour les items relatifs au « prélèvement de l'information » et 32 % pour ceux relevant de « l'organisation de l'information » et à seulement 27,6 % pour ceux de la compétence « interpréter ». Ces élèves ont des connaissances très fragmentaires et restreintes. Ils sont capables de rechercher et prélever quelques informations très explicites sur des supports simples (chiffres dans un tableau, frontière ou ville sur une carte, courbes statistiques très différenciées).

Les élèves du groupe 2 (28,1 % du total) ont un pourcentage moyen de réussite à l'ensemble des items de 47,4 %. Ils réussissent en moyenne un item sur deux relevant d'« identifier les informations ». Ils savent travailler sur les supports documentaires de la discipline à partir desquels :

- ils sont capables de reconnaître et de nommer des pays, des villes, des mers et océans pré indiqués (fléchés ou pointés) sur des représentations cartographiques différentes (cartes conventionnelles, projection polaire et anamorphose) à l'échelle mondiale et à l'échelle européenne. En ce sens, on peut dire qu'ils ont mémorisé les configurations territoriales ;
- ils peuvent reconnaître, nommer et replacer dans un ordre chronologique quelques figures historiques (grands empereurs de l'Antiquité, du Moyen-Âge, du XVIII^{ème} et du XIX^{ème} siècle par exemple) à partir de documents iconographiques ;
- ils utilisent le langage de la carte : lecture de légende, de figurés simples, de signes conventionnels ou interprétatifs pour prélever des informations explicites ;
- ils restituent des connaissances correspondant aux repères du programme de troisième et ceux attendus et listés pour le diplôme national du brevet (DNB)¹. Ils ont mémorisé de manière parcellaire quelques

NOTE

1. Une épreuve du DNB porte sur la connaissance de repères chronologiques et spatiaux dont la liste figure dans les programmes de troisième (*Bulletin Officiel*, n° 10 du 15 octobre 1998).

connaissances sur les programmes antérieurs ;

– ils définissent le lexique spécifique aux disciplines (lexique économique, juridique, historique simple), mais celui-ci reste limité ;

– ils listent des éléments se rattachant à des faits (récits se rattachant à la Bible, fondateurs des religions monothéistes, contextes d'événements de la Résistance et de la Guerre froide, par exemple) et sont capables de hiérarchisation et de classement simple (réussite à 44,3 % pour les items relevant de la compétence traiter l'information).

Les élèves de ce groupe ne maîtrisent pas suffisamment la compétence « interpréter » (réussie à 43,1 %). Ils ne donnent du sens qu'à quelques documents iconographiques spécifiques (caricature, photographie à usage de propagande) ou textuels (sens d'un court extrait d'un récit de déporté par exemple). L'écrit constitue visiblement un frein (46 % de non-réponses aux questions ouvertes et faible réussite).

Le groupe 3 (30 % des élèves) réussit en moyenne deux items sur trois (63,7 %). Il a une meilleure maîtrise des compétences des groupes précédents. Ainsi le pourcentage moyen de réussite est de 67 % pour la compétence « identifier ».

Les repères (localisation, dates, événements) attendus pour le DNB sont acquis, mais également des repères se rattachant aux programmes antérieurs de sixième et cinquième (textes sacrés des religions monothéistes par exemple). Ils ont une perception du temps à différentes échelles (ils situent des personnages, des événements, des faits de civilisation de l'Occident chrétien dans les grandes périodes de l'histoire), ils datent de manière précise et localisent des faits historiques (Hégire,

Constantinople, Grandes découvertes, Louis XIV par exemple).

De plus, ils répondent aux questions ouvertes lorsqu'elles appellent des réponses brèves leur demandant de mobiliser ce qu'ils ont mémorisé (date, nom propre, devise de la France, signification des fêtes nationales par exemple) ou de recopier des arguments explicitement désignés.

Ces élèves réussissent en moyenne 64 % des items évaluant la maîtrise du vocabulaire (« décrire »). Celui-ci est plus précis et élargi à des champs variés et plus nombreux (juridique, élections en France, laïcité), économique, politique (nazisme, communisme), et religieux (religion d'État) et ils ont accès au vocabulaire conceptuel (libéralisme). Ils maîtrisent certains concepts fondamentaux du programme de troisième en histoire, géographie et éducation civique.

La lecture de textes longs ou de propositions de réponses longues et complexes n'est plus un frein à la réalisation des tâches demandées.

Les items relatifs à la compétence « traiter l'information » sont réussis à 58 %. Les élèves gèrent des documents et des tâches complexes (textes courts et longs, organigrammes, textes et photographies, croisement de plusieurs informations, confrontation de documents...) à partir desquels ils sont capables de hiérarchisations complexes et de mise en relation (notamment causes/conséquences). Ils donnent du sens et pèsent la portée de documents.

Les items relatifs à la compétence « interpréter » sont réussis à 60,2 %. Les élèves parviennent à généraliser : intituler, dégager l'idée à partir d'une description ou d'une argumentation (texte de 15 lignes sur les enjeux spatiaux, économiques et politiques d'un aménagement de territoire par

exemple). Ils passent du particulier au général et inversement ; ils justifient en citant le texte (en question ouverte). Ils perçoivent la différence de points de vue sur un événement, une période, un fait de société, à partir de documents (textes, affiches, caricatures...)

Un item sur deux en moyenne de la composante « réaliser » est réussi : Les élèves donnent du sens à des figurés simples et usuels de cartes schématisées et de croquis et perçoivent le sens de la proportionnalité. Ils reconnaissent des figurés de légendes et de cartes à partir de photographies de paysages littoraux ou agricoles par exemple.

Les élèves du groupe 4 représentent 16,9 % de l'ensemble. Leur pourcentage moyen de réussite à l'évaluation passe à 77,3 %. En plus d'une plus grande maîtrise des compétences des groupes précédents, leurs connaissances sont approfondies et installées sur les temps longs et les temps courts. Ils peuvent mobiliser des connaissances qui, pourtant, n'ont pas été réactivées par les programmes des classes antérieures.

Le vocabulaire (items réussis à 77 %) est maîtrisé : aussi bien le vocabulaire spécifique (sigles compris) que le vocabulaire conceptuel (plus élargi et plus difficile) auquel ils donnent du sens (république, démocratie, les différentes formes de pouvoir, les principes de la République française, les mots clés du XIX^e siècle...).

Les trois quarts des items « interpréter » sont réussis. Les élèves sont capables de donner les idées essentielles de documents variés, longs et complexes (textes longs, affiches, textes de lois, chronologie détaillée, tableaux complexes...) en les comparant. Leurs connaissances et leur capacité à synthétiser leur permettent d'accéder rapidement à l'implicite.

Ces élèves commencent à appréhender la démarche des disciplines et à l'intégrer dans leur analyse (critique des sources, des documents, utilisation du contexte historique, utilisation du para-texte (auteur, titre, date), sens et utilité des symboles graphiques...), ce qui leur permet de justifier, argumenter, critiquer des affirmations.

Dans les questions ouvertes, ils sont capables de trouver plusieurs arguments demandés et parfois sans guidance, en utilisant les termes et notions appropriés. Ils peuvent expliquer et justifier une représentation graphique. Les items de la composante « réaliser » sont réussis en moyenne à 70 %. Les élèves de ce groupe construisent des légendes et des cartes en utilisant les figurés et leurs proportions appropriées. Ils peuvent passer d'un langage à l'autre, tout en conservant le sens : réaliser la légende d'un croquis issu d'une photographie et d'un texte, compléter une carte à partir des informations d'un texte.

Les élèves du groupe 5 (10 %) ont un taux moyen de réussite à l'ensemble de l'évaluation de 87,5 %.

Ils réussissent 90 % des items « identifier ». Par rapport aux groupes précédents, ils ont en plus des connaissances très précises concernant par exemple les datations. Les savoirs sont nettement installés depuis la sixième.

De même ils réussissent 85,4 % des items « interpréter ». Par rapport aux groupes précédents, ces élèves vont directement au sens, à l'analyse critique.

De plus, l'essentiel de l'argumentation passe par la phase rédactionnelle : c'est le seul groupe qui répond de manière réussie à la quasi-totalité des questions ouvertes. On peut dire de ces élèves qu'ils respectent les consignes, vont au bout du travail demandé : réponses à tous les arguments

demandés, conceptualisation, vocabulaire précis, réponses construites.

CE QUE CES RÉSULTATS RÉVÈLENT

Au regard des attendus de fin de collège

En regard des contenus des programmes, de leurs préconisations et des attendus de fin de collège² on peut dire que, à la fin du collège : la grande majorité des élèves (85 %, c'est-à-dire à partir du groupe 2) est habituée à travailler avec le matériel documentaire de la discipline. Ils ont acquis au moins la lecture documentaire. Cette évaluation est révélatrice des pratiques d'enseignement : les résultats des élèves témoignent de leur familiarité avec ce matériel sous toutes ses formes, notamment les documents iconographiques.

Cependant les élèves du groupe 2 ne se limitent qu'à la lecture séquentielle de ces documents qui leur permet de les décoder, de réactiver des connaissances et de réussir certaines compétences comme le prélèvement ou le classement d'informations simples. Ils ne vont pas au-delà. Pour ces élèves, connaissances et compétences restent ponctuelles et dans l'explicite. Il est à noter que l'écrit rebute les élèves de ce groupe (50 % de non-réponses aux questions ouvertes).

À partir du groupe 3, plus de la moitié des élèves (soit 57 %), maîtrisant mieux le matériel documentaire, ont la capacité à utiliser des documents plus complexes, notamment des textes plus longs (lecture cursive et analytique – programme de troisième). Ils maîtrisent essentiellement les repères liés au programme de troisième et ceux attendus au DNB (sens de la durée et de l'espace – programme de troisième).

Ils donnent du sens à leurs connaissances. Leur vocabulaire est plus élargi, plus précis. Ils sont capables de hiérarchisations complexes et de mises en relation de supports différents. Ils parviennent à généraliser, commencent à justifier et perçoivent la différence de points de vue. Il semble-là que se situe un « noyau dur » de connaissances et de compétences disciplinaires sans doute bien consolidées au moment de la passation proche des épreuves du brevet des collèges. C'est dire que l'année de troisième joue pleinement son rôle dans la consolidation des connaissances et dans la préparation à l'examen.

27 % des élèves (groupes 4 et 5) font preuve de connaissances installées portant sur l'ensemble des programmes du collège. Ils maîtrisent en outre les concepts difficiles et sont désormais dans l'implicite. Leur aptitude à la lecture et à l'appréhension de la démarche de la discipline leur permet de réaliser toutes les situations en allant jusqu'au bout du travail, de répondre à toutes les questions en respectant toutes les consignes. Ce sont eux qui réalisent les productions écrites. C'est dans la réussite des compétences en argumentation et en rédaction que se dessine une ligne de démarcation où ne réussissent que les meilleurs élèves (mais le programme de troisième précise qu'il s'agit d'un apprentissage). La production écrite – il s'agit sans doute là d'une évidence – ne leur fait pas peur et ils affronteront avec aisance l'épreuve du paragraphe argumenté du DNB.

NOTE

² Accompagnement des programmes (Bulletin Officiel n° 25 du 20 juin 1996).

Au regard du passage à la production écrite

Demander aux élèves des productions spontanées écrites paraissait nécessaire pour trois raisons : d'une part, ce format de questions dites « questions ouvertes » est usuel dans notre système éducatif ; d'autre part, il permet de vérifier que les élèves ont mémorisé des repères, des faits de civilisations par exemple... et qu'il les mobilisent à bon escient sans indices venant aider à raviver la mémoire (comme le format QCM peut le faire) ; enfin, évaluer des productions écrites plus longues correspond, en histoire-géographie, à la formation des élèves qui « depuis la sixième ont été entraînés à élaborer de manière autonome quelques phrases simples » et « en classe de troisième [ils] sont progressivement capables de rédiger un paragraphe cohérent de quinze à vingt lignes » (programme de troisième) et à « utiliser d'autres langages » (construction de cartes et légendes) à des fins d'analyse, de synthèse ou d'argumentation, entre autres, pour

lesquelles l'élève doit structurer ses connaissances, organiser son message pour le communiquer.

Lors de la phase expérimentale, 36,7 % des élèves n'ont pas renseigné les questions ouvertes. L'épreuve finale comprenait 58 items regroupant 91 prélèvements d'informations. Le taux de non-réponses reste toujours élevé : 34,4 %, (contre 7 % de non-réponses aux questions fermées) sans différence significative entre les questions appelant des réponses brèves ou longues. En moyenne, les questions ouvertes sont moins bien réussies (39,8 %) que les QCM (62 %). Pour chaque groupe d'élèves, il y a un écart de 20 points (graphique 3)

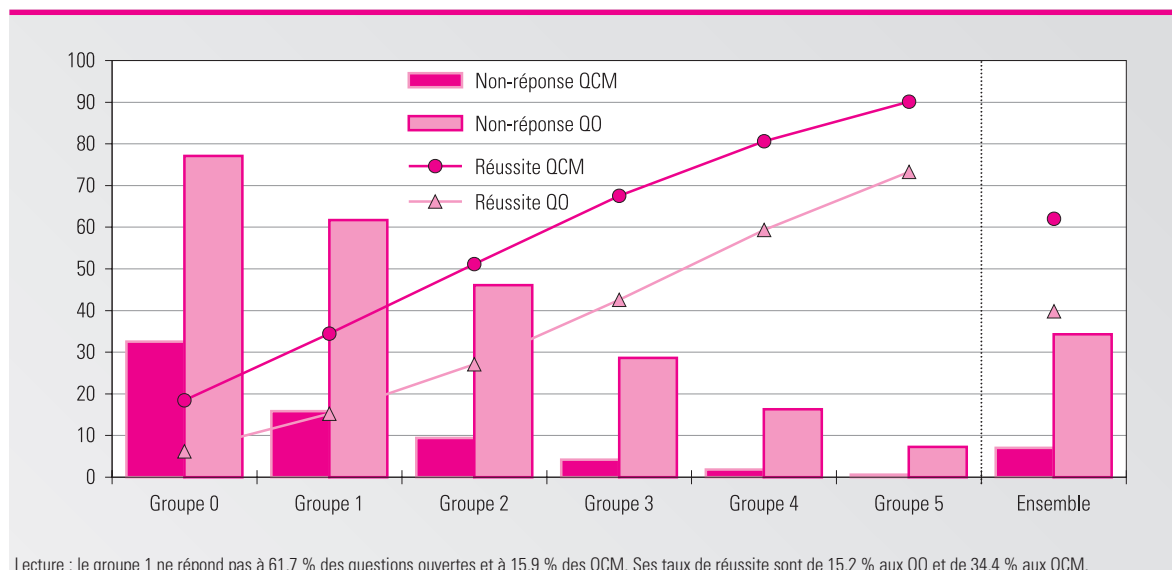
Les élèves faibles (groupe 2) ne sont capables d'écrire une réponse brève qu'avec l'aide d'un support documentaire (texte, photo...). La mémoire sans réactivation est mobilisée avec succès à partir du groupe 3.

Les items appelant des réponses construites et souvent proposées avec une guidance plus ou moins prégnante (d'une demande d'une simple phrase à

la rédaction d'un à trois paragraphes généralement clairement identifiés) visaient essentiellement la compétence « interpréter » (généraliser, argumenter, réaliser). Ces items ne sont réussis à plus de 50 % que par les élèves appartenant aux groupes 4 et 5, (contre seulement 39 % par ceux du groupe 3). On a donc affaire à un haut niveau de compétence : trouver et écrire des arguments en autonomie dans des documents ou encore trouver ses propres arguments. On peut dire qu'en ce qui concerne les productions écrites, les élèves de ces deux groupes sont allés jusqu'au bout des tâches demandées.

Ce taux de non-réponses et de moindre réussite aux questions ouvertes est surprenant pour des élèves qui devraient être familiarisés avec cette forme de questionnement. Est-ce dû au faible enjeu de cette évaluation anonyme pour les élèves ? Le QCM apparaît sans doute comme une tâche plus attractive tandis que la réponse écrite demande un effort rédactionnel, aussi court soit-il. Cet effort qui nécessite une certaine maîtrise de la langue

Graphique 3 – Taux de non-réponse et de réussite par groupe selon le type de questions



Source : MEN-DEPP

française à travers des productions longues, correspond en outre à l'évaluation de composantes disciplinaires plus complexes : « généraliser », « synthétiser » « argumenter ». Laquelle de ces composantes fait écran à l'autre ?

La question de l'attitude des élèves face à l'écrit au cours de cette évaluation est à mettre en regard de l'opinion des enseignants interrogés dans le cadre d'une enquête sur leurs pratiques³, concernant les difficultés à enseigner l'écriture d'un paragraphe et mérite réflexion : qu'est-ce qu'apprendre à rédiger et quel temps accorder à cet apprentissage, qu'est-ce qu'évaluer l'écrit et comment évaluer l'écrit, dans des disciplines autres que le français ? Qu'est-ce qui inhibe l'élève et quel levier envisager pour la formation et l'apprentissage ? Cette évaluation pointe la question de la place de l'écrit en une fin de cycle qui prépare à l'épreuve du paragraphe argumenté du brevet des collèges et à l'entrée au lycée où, surtout dans les formations générales, on fera essentiellement appel à cette forme d'expression. Or, au fil du collège, les élèves sont formés à l'écrit, pour passer de la rédaction de phrase simple (sixième) à des phrases plus élaborées (cycle central) et à la rédaction du paragraphe argumenté (troisième).

Au regard de pratiques évaluatives

Les enseignants concepteurs étaient habitués aux évaluations conduites en classe aux caractéristiques spécifiques : connaissance des progrès d'élèves connus, attente de l'enseignant, note attribuée sur 20, interrogation sur une partie du programme. Ils se sont trouvés confrontés dans le cadre de

l'évaluation-bilan nationale à une autre échelle : un constat objectif des acquis de l'ensemble des élèves en France quels que soient les contextes pédagogiques (établissements publics – dont les ZEP – et privés) et sur l'ensemble des programmes de collège. Il leur a fallu répondre au cadre méthodologique posé par la commande et se questionner sur les pratiques évaluatives quotidiennes.

Comment évaluer l'ensemble des acquis du collège ?

Évaluer les acquis des élèves de la sixième à la troisième, c'est évaluer en histoire, géographie et éducation civique, un corpus de connaissances quantitativement très important. Le regroupement dans des thèmes, les points forts dégagés des programmes, le nombre conséquent de situations et d'items permettent d'avoir suffisamment d'informations pour que le traitement statistique s'attache, dans un premier temps, non pas à pas à telle ou telle connaissance ponctuelle, mais à un ensemble. Par exemple les élèves ont été questionnés sur un nombre suffisant de personnages ou de faits historiques pour estimer un seuil de connaissances maîtrisées sur des périodes historiques et des espaces géographiques.

Comment cibler les compétences visées ?

En classe, l'enseignant fait souvent une appréciation globale du travail de l'élève où l'intuitif peut jouer un rôle d'autant qu'il connaît ses élèves (leur niveau, leurs progrès) et qu'il vient de faire cours en s'étant fixé des objectifs. Un bilan des acquis sur un corpus conséquent de connaissances et de compétences nécessite de « décomposer » tout ce qui participe à la réponse

de l'élève, l'environnement de la question et l'élève lui-même. Toute une réflexion a été conduite pour bien cibler, par les questions posées, la compétence visée, être vigilant pour évacuer (ou minimiser) tout ce qui pouvait parasiter l'objectif visé :

- le choix du vocabulaire pour la rédaction des consignes,
- les supports documentaires utilisés couramment en histoire-géographie : évalue-t-on la maîtrise de la lecture documentaire ou la compétence visée à travers le document ?
- la tâche demandée : évalue-t-on la tâche demandée ou la compétence visée, sachant qu'une compétence est bien maîtrisée si elle est mobilisée et dominée dans des situations différentes. Quel est le poids de celles-ci dans ce que l'on évalue ?

Ces interrogations ont été particulièrement sensibles pour les questions ouvertes (pourtant questionnement plus proche des pratiques de classes que les QCM). Quelle doit être la tolérance admise pour évaluer la localisation des villes françaises quand Lyon est écrit Lion, Lille, Lile ou l'île etc. ? L'évaluation de cette compétence serait-elle la même à l'oral où l'on ne vérifie pas l'orthographe ?

De plus, une même production écrite d'élève de deux lignes peut donner des informations pour quatre composantes

NOTE

3. Voir l'article « Les pratiques d'enseignement en histoire-géographie et éducation civique au collège » dans ce numéro.

Image de la discipline et pratiques d'enseignement en histoire-géographie et éducation civique au collège, Les dossiers, n° 183, MEN-DEPP, mars 2007.

« Image de la discipline et pratiques d'enseignement en histoire-géographie et éducation civique au collège », *Note d'Information* 07.27, MEN-DEPP, mai 2007.

différentes. Exemple : « autrefois l'hiver en montagne était considéré comme une saison morte : la situation est-elle toujours la même aujourd'hui ? justifiez votre réponse en utilisant les documents » (une photographie et un texte décrivant des activités en montagne avant et après 1950). Composante généraliser (sens de la réponse) ; composante justifier (arguments valables quels qu'ils soient) ; composante sélectionner (arguments valables pris uniquement dans les deux documents) ; composante rédactionnelle (critères retenus pour l'expression française).

À cet égard, deux débats peuvent alors s'engager :

- pour évaluer les acquis en histoire, géographie ou éducation civique, a-t-on besoin de passer par la production écrite des élèves ? ou : quand on passe par la production écrite des élèves, qu'évalue-t-on exactement ? la maîtrise de la langue : orthographe, syntaxe, utilisation des connecteurs logiques, ou la maîtrise de compétences spécifiques à la discipline : vocabulaire spécifique, explication d'une situation historique, critique d'un document ?
- y-a-t-il des compétences propres à l'histoire-géographie, éducation civique ou seulement des compétences transversales (l'histoire-géographie participant de la maîtrise de l'écrit) ?

Comment élaborer des questions... et des réponses ?

Le questionnement devait se faire pour une grande part sous forme de QCM. Cette pratique est peu développée en classe. Dans ce format de question, l'élève a moins d'autonomie et sa mémoire est réactivée par les propositions de réponse. En revanche, il n'est pas confronté à la production écrite. Il n'en reste pas moins que les tâches de lecture sont plus diverses et

plus nombreuses (des consignes, des différentes propositions de réponse...) et que la compréhension du vocabulaire est plus exigeante (car il y a plus de nuances lexicales dans les différentes propositions de réponse qui ne se contentent pas d'à peu près). En fait, la « fabrication » des propositions de réponses (et, parmi elles, la bonne réponse attendue par les enseignants) ne fut pas toujours aisée notamment pour obtenir un consensus sur certaines définitions qui semblaient à chacun au départ évidentes :

- rigueur des mots employés et compréhensibles par les élèves (dans tout type d'établissement) ;
- rédaction des autres propositions que celles attendues (les distracteurs) à partir de l'expérience des concepteurs (erreurs récurrentes des élèves, représentation).

L'enseignant se met en situation de production de réponses (justes, approchantes ou fausses) et non pas seulement de questions. Il se trouve alors confronté à un effort de rigueur scientifique et de vulgarisation (consensus pour tous les élèves).

Ce type d'évaluation est donc une contribution parmi d'autres avec ses intérêts et ses limites à la réflexion sur l'évaluation des élèves. Elle enrichit les autres formes d'évaluation et s'enrichit de celles-ci.

Au regard de l'image que les élèves ont de la discipline

Les questionnaires cognitifs des élèves étaient accompagnés de questionnaires de contexte permettant de cerner l'environnement familial et scolaire des élèves ainsi que leur perception de la discipline ou de leur établissement. Quelques aspects de leurs réponses méritent attention.

Des performances qui varient selon l'importance accordée à la discipline...

Les élèves étaient interrogés sur l'importance qu'ils accordent à la discipline et il en ressort que plus les collégiens estiment qu'elle est importante, meilleurs sont leurs résultats. Ainsi, les élèves qui répondent que, selon eux, l'histoire-géographie, l'éducation civique est « très importante » ou « importante » sont surreprésentés dans les groupes de haut niveau, c'est-à-dire les groupes 4 et 5 et, à l'inverse, ceux qui considèrent cette matière comme n'ayant pas ou peu d'importance sont surreprésentés dans les groupes faibles (0 et 1).

... le temps de travail personnel...

Les performances des élèves sont également très différenciées selon le temps qu'ils passent à travailler la discipline hors de la classe. Le travail hors de la classe semble avoir un impact très positif sur les résultats. Ainsi, 16,4 % des élèves qui disent travailler entre 1 et 2 heures par semaine chez eux font partie du groupe 5. Au contraire, les élèves déclarant peu travailler (moins d'une demi-heure par semaine) sont moins nombreux que l'ensemble à appartenir aux meilleurs groupes (8,8 % disant travailler moins de un quart d'heure sont dans le meilleur groupe) et sont surreprésentés dans les groupes faibles (22,3 % des élèves travaillant moins de un quart d'heure se trouvent dans les groupes 0 et 1). Ce résultat n'est pas habituellement observé dans d'autres disciplines pour lesquelles une importante proportion d'élèves travaillant peu en dehors de la classe fait pourtant partie des groupes de hauts niveaux. Il semble que pour avoir de

bons résultats en histoire, géographie et éducation civique, un important travail personnel est indispensable pour pouvoir acquérir le corpus de connaissances nécessaire à la réussite des tâches demandées.

... et les habitudes culturelles en lien avec la discipline...

Plus les élèves sont en contact avec la discipline par le biais de livres et de revues, meilleurs sont leurs résultats. Ainsi, 21,7 % de ceux qui répondent s'intéresser « souvent » ou « très souvent » à la matière en lisant, appartiennent au groupe 5 contre 6,9 % de ceux qui répondent « jamais » ou « rarement ». Au contraire, ces derniers sont surreprésentés dans les groupes faibles. La fréquence de la lecture de livres ou de revues influence davantage le score en histoire qu'en géographie et en éducation civique.

De même, les visites fréquentes de musées, de sites historiques ou de lieux géographiques ont une importante influence positive sur les résultats des élèves. Ainsi, les collégiens qui répondent « très souvent » ou « souvent » sont surreprésentés dans les groupes de haut niveau (45,8 % dans les groupes 4 et 5) contrairement aux élèves qui

choisissent les modalités « jamais » ou « rarement » qui sont moins de 25 % à appartenir à ces groupes. L'influence de ces visites est plus importante sur le score en histoire. En revanche, l'influence des émissions de télévision ou de films, de même que celles des cédérom et d'Internet est moins nette sur le score des élèves et sans différence significative entre les disciplines.



Les élèves dessinent une image de la discipline plutôt attractive, nécessitant un travail hors de la classe et un enrichissement personnel qui influent sur leurs résultats.

Cette évaluation démontre qu'une grande majorité d'élèves a accès à la lecture documentaire de tous les supports utilisés dans la discipline, conformément aux préconisations officielles et reflétant les pratiques pédagogiques. Ces résultats mettent en évidence la prégnance des documents en histoire, géographie, éducation civique ce qui invite à questionner le statut didactique et pédagogique qu'on leur accorde. Quelle stratégie pédagogique, en effet, mettre en place pour que 30 % des élèves puissent

développer une plus grande autonomie et des compétences plus complexes à l'égard de ces supports ? Ils rejoindraient alors le tiers des élèves qui ont des connaissances et compétences plus élargies. Pour l'instant, 27 % des élèves vont au bout du travail demandé (rédactionnel compris) correspondant à tous les attendus de fin de collège. D'où l'intérêt de comparaisons temporelles que permettent les évaluations-bilans. L'effet année en cours (troisième) et la proximité des épreuves du DNB favorisent pour la moitié des élèves l'acquisition d'un bagage de connaissances et de compétences consolidé dont l'examen pourrait aider à la réflexion sur le socle commun. Cette évaluation pose aussi la question de la place et du poids de l'écrit dans la discipline : taux important de non-réponses, pourcentages de réussite aux questions ouvertes inférieurs aux QCM, mais aussi difficultés à évaluer objectivement les productions écrites des élèves. Elle permet de questionner les pratiques évaluatives habituelles (objectifs visés, adaptation du questionnement, critères d'évaluation...). Autant de questions qui mériteraient d'être approfondies et utilisées avec profit par les décideurs. ■