

PISA

Franck Salles(*)

PISA est sans doute l'enquête internationale sur les acquis des élèves la plus médiatisée. Et par conséquent la plus critiquée. Le texte suivant propose une description du déroulement de la conception du test. Ce point de vue « de l'intérieur » permettra peut-être au lecteur de mieux connaître l'enquête et d'envisager ainsi avec plus de recul l'utilisation, les annonces et les critiques faites à son sujet.

L'enquête PISA : une vue d'ensemble

PISA (*Program for international student assessment* ou en français : *Programme international pour le suivi des acquis des élèves*) évalue les aptitudes des élèves de 15 ans dans les trois domaines de « literacy » (en anglais dans le texte) : culture mathématique, culture scientifique et compréhension de l'écrit [1]. À lui seul le critère d'âge donne au PISA un caractère original. Les autres évaluations internationales, ou nationales en France, évaluent en effet un niveau scolaire. À titre d'exemple, TIMSS [2] évalue en fin de quatrième et en fin de seconde (la France ne participe plus à cette étude depuis 1995). CEDRE [3] (bilans nationaux collège et primaire de la DEPP) évalue en fin de troisième et fin de CM2, pour ne citer que deux exemples.

Ce choix correspond à peu près à la fin de la scolarité obligatoire dans la plupart des pays participants, pas en France cependant où il est de 16 ans. Il implique une exception française dans l'enquête internationale : en France, les élèves de 15 ans se répartissent dans deux types d'établissements scolaires différents : le collège et le lycée (voir table 1). C'est une conséquence de la politique de redoublement en France : environ 35% des élèves de 15 ans sont « en retard » en France en 2009 et sont donc encore en collège. C'est le taux le plus élevé des pays participant au PISA [4].

<i>Table 1 : Répartition des élèves de 15 ans (nés en 1993) ayant participé à l'évaluation PISA en 2009</i>		
	Classe fréquentée	Répartition en %
En avance	Première générale et technologique	2,5
À l'heure	Seconde générale et technologique	51,4
	Seconde professionnelle	9,2
En retard	Troisième	31,9
	Quatrième	3,6
	Autre ou inconnu	1,4
<i>Lecture: 51,4% des élèves de l'échantillon sont en Seconde GT.</i>		
<i>Source: MEN-DEPP/OCDE</i>		

(*) Professeur de mathématiques, Chargé d'études sur PISA à la DEPP.
franck.salles@education.gouv.fr

L'enquête PISA : les différents acteurs

L'OCDE pilote l'enquête, procède à l'analyse des données et communique les résultats lors d'une conférence de presse très médiatisée [5]. Elle commande la construction et l'organisation du test à un consortium d'établissements divers. Ce consortium est international (Australie, États-Unis, Belgique, Allemagne, Japon, Norvège, Luxembourg). Les établissements qui le composent sont parfois issus de la sphère publique (universités, « non profit groups » (en anglais dans le texte)), semi publique ou privée. Ce consortium agit sous l'expertise de groupes d'experts dans chaque domaine. Ces experts sont généralement issus du monde universitaire ou de l'évaluation.

Chaque pays participant possède un centre national PISA. En France, le centre national PISA est au bureau de l'évaluation des élèves de la DEPP (*Direction de l'évaluation, de la prospective et de la performance* au Ministère de l'Éducation Nationale). Les centres nationaux interviennent localement et sont consultés tout au long du processus de construction ou de mise à jour de l'enquête. La mission des centres nationaux est d'administrer le test dans les conditions standard imposées par le consortium international. Ils comprennent dans leur équipe des statisticiens, des administrateurs et des enseignants chargés d'études dans les domaines. En France, à la DEPP, les chargés d'études coordonnent le travail de groupes d'enseignants, formateurs et inspecteurs pour participer aux travaux de conceptions d'exercices PISA, à l'analyse critique du développement du cadre de l'évaluation et à l'analyse des résultats.

L'enquête PISA : description d'un cycle

L'évaluation a lieu tous les trois ans depuis 2000. Pour chaque cycle et à tour de rôle, un des trois domaines est majeur, c'est-à-dire que les deux tiers des exercices du test relèvent de ce domaine, le reste relevant des deux autres domaines⁽¹⁾. Ainsi tous les 9 ans, une comparaison peut se faire sur un domaine majeur à partir d'un nombre significatif d'exercices. Cette comparaison a été possible pour la première fois entre les cycles PISA 2000 et 2009 en compréhension de l'écrit [6]. En 2012, le domaine de culture mathématique sera domaine majeur comme il l'avait été en 2003.

Les étapes d'un cycle PISA

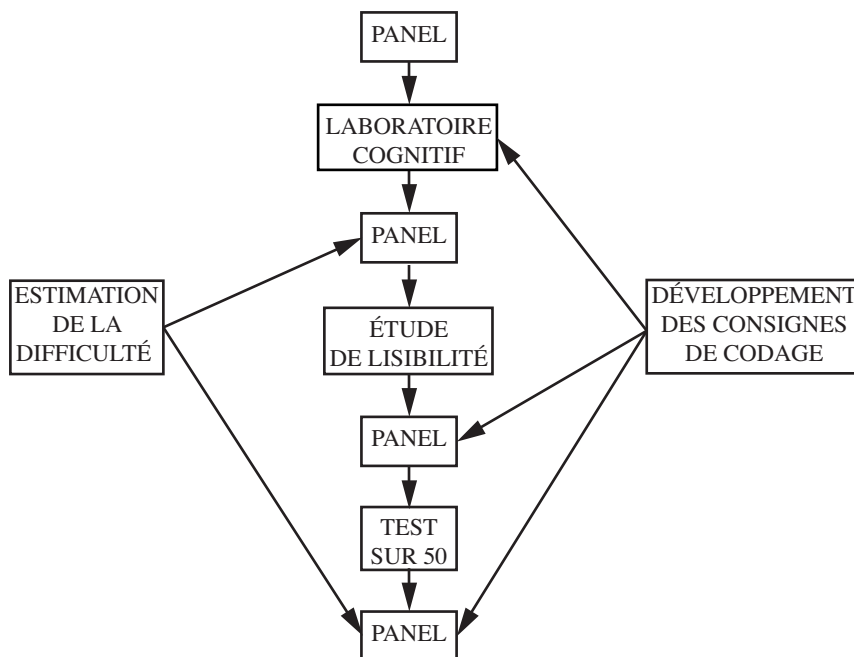
1. L'OCDE commande au consortium la conception d'un cadre théorique d'évaluation. Ce cadre définit notamment la notion de « literacy » qui est évaluée chez les élèves de 15 ans dans le monde. ACER (*Australian Council of Educational Research*) est dans le consortium le principal établissement qui construit et rédige ce cadre. Le cadre d'évaluation est publié dans les deux langues officielles de l'OCDE, Anglais et Français, et est téléchargeable gratuitement sur le site de l'OCDE [1]. Les centres nationaux sont seulement consultés sur le développement du cadre

(1) Les majeurs : 2000, compréhension de l'écrit ; 2003, culture mathématique ; 2006 : culture scientifique ; 2009, compréhension de l'écrit.

d'évaluation. Le cadre est mis à jour tous les trois ans mais aucun changement drastique ne peut intervenir au risque d'interférer sur les comparaisons d'un cycle à l'autre. La définition de la culture mathématique a par exemple évolué depuis 2003 et le cadre 2012 apporte une définition légèrement différente.

2. Une fois le cadre établi, vient une phase de construction d'exercices pour le test lui-même. Le consortium, notamment ACER(Australie), ainsi que les centres nationaux engagent ces travaux de conception. En 2003 par exemple, la moitié des exercices de mathématiques du test étaient issus des productions des pays. Pour 2012 en revanche, ce pourcentage sera moindre (environ 20%).

Le bureau de l'évaluation des élèves de la DEPP a procédé à de tels travaux pendant l'année 2010 pour la culture mathématique. Un groupe de travail (appelé **panel**) a été mis en place comprenant enseignants, inspecteurs pédagogiques régionaux, inspecteurs généraux, formateurs, chercheurs et chargés d'études. La conception d'un exercice suit la procédure suivante :



Pour lire cet organigramme ;

- la colonne centrale donne les phases successives de l'élaboration ;
- les flèches de gauche indiquent les phases où se fait l'estimation de la difficulté ;
- celles de droite les phases du développement des consignes de codage.

Pour des détails sur « laboratoire cognitif » et « consignes de codage », voir les annexes 3 et 4.

L'exercice est testé sur une cinquantaine d'élèves (« test sur 50 » dans l'organigramme ci-dessus) et à l'issue de ce processus, l'item est soumis au consortium international qui vérifie la pertinence de l'exercice compte tenu du cadre et opère une première sélection.

3. Le consortium soumet ensuite (en anglais) l'ensemble des exercices conçus aux 66 centres nationaux pour repérer des biais culturels ou prévoir des impossibilités de traduction éventuelles. Une seconde sélection s'opère ainsi.

4. Les exercices sont ensuite traduits. En 2003, tout le matériel de test a été traduit en 33 langues. Chaque pays PISA (sauf les pays francophones et anglophones) prend en charge la traduction à partir des deux langues sources (anglais et français) mais le consortium, notamment Capstan (Belgique), participe aux processus de vérification. Ces versions sont ensuite adaptées : les versions belge et française sont par exemple légèrement différentes pour ne pas désavantager ou avantager les élèves d'un de ces pays.

On peut comprendre l'importance cruciale de cette étape dans la standardisation des exercices. La tâche de l'élève doit en effet être la même quelle que soit sa langue. Or le registre langagier est très présent dans les exercices PISA de culture mathématique (voir annexe 1).

En 2009, un item PISA (stimulus et question) contient en moyenne 100 mots (médiane 72 mots). Le plus long en contient près de 400 ! Une approche qualitative montre que l'élève doit mettre en œuvre ses compétences de lecture pour s'approprier une situation, inférer sur le vocabulaire pour orienter son activité mathématique. Dans l'item « le menuisier » (voir annexe 2), un seul mot de l'énoncé engage l'élève à mobiliser ses connaissances et savoir-faire autour de la notion de périmètre : le mot « bordure ». Si le sens de ce mot est mal perçu, il y a fort à parier que l'activité mathématique ne peut pas suivre...

5. Les exercices arrivés à ce stade de la conception sont ensuite soumis à un test grandeur nature (pour le PISA 2012, 172 questions de culture mathématique sont arrivées à ce stade). Il s'agit d'un essai de terrain qui a lieu dans tous les pays PISA, un an avant le test final (dernier essai de terrain PISA pour les exercices de culture mathématique : mai 2011). L'essai de terrain porte sur un échantillon d'élèves réduit de moitié environ par rapport au test final.

Il s'agit de déterminer la qualité des exercices en termes psychométriques, autrement dit tenter de répondre à la question suivante pour chacun d'entre eux : l'exercice est-il réussi par des élèves d'aptitude supérieure à ceux qui échouent ? Une ultime sélection est ainsi faite, à partir des données statistiques issues de l'essai de terrain. Pour le PISA 2012, une quarantaine de nouvelles questions de culture mathématique (en plus des 36 questions issues des années précédentes) vont ainsi être choisies sur les 172 propositions initiales.

6. Pour chaque pays, l'échantillon est tiré par le consortium, notamment WESTAT (États-Unis), en deux temps : par établissements, puis par élèves. Les centres nationaux fournissent au consortium la liste exhaustive des établissements scolaires

accueillant au moins un élève de 15 ans. Puis dans chaque établissement de l'échantillon, un échantillon des élèves d'âge requis est aléatoirement tiré. Pour le test final en 2009, l'échantillon français comportait 150 établissements (collèges, lycées généraux et techniques, lycées agricoles) et 32 élèves par établissement.

7. À l'issue de la passation du test, les cahiers des élèves sont corrigés par les centres nationaux. La méthode de correction dépend du format de la réponse. En cas de QCM, les réponses sont enregistrées automatiquement, à l'aide d'un logiciel de lecture optique. En cas de réponse ouverte, la correction est assurée par des professionnels de l'enseignement de la matière dans chaque pays. Les consignes de correction sont fortement standardisées et des formations internationales sont organisées en amont pour tenter d'harmoniser ces consignes.

Là encore il est aisé de comprendre l'importance de cette étape dans la standardisation du test, des centaines de correcteurs de par le monde doivent en théorie être capables de donner le même crédit à la même réponse d'élève. Lorsqu'on connaît les résultats d'études docimologiques [7], on sait qu'il s'agit d'un vœu pieux ! Pourtant le consortium met tous les atouts de son côté en développant, à l'aide des concepteurs d'exercices, des consignes de correction très précises pour ces questions ouvertes. De manière générale, le correcteur doit donner un crédit à la réponse qui manifeste que ce qui est attendu est acquis. Dans certains cas, il a la possibilité de donner un crédit partiel lorsque l'activité de l'élève a été incomplète. Dans d'autres cas, il peut coder deux types de crédits équivalents pour des méthodes de résolutions différentes. (Voir annexe 4 pour exemple).

8. Enfin l'ensemble des données ainsi récoltées est analysé par l'OCDE. Elle publie ces conclusions généralement 18 mois après la passation finale. La prochaine passation du PISA aura lieu dans le monde en mai 2012 et les premiers résultats seront rendus publics en décembre 2013. Les centres nationaux ont aussi l'opportunité d'analyser les résultats des élèves de leur pays, mais tout résultat est sous « embargo » jusqu'à la conférence de presse officielle de l'OCDE.

L'évaluation PISA : à la mesure de l'objectif

L'ambition et la taille du PISA en font l'enquête internationale sur les acquis des élèves la plus médiatisée et sans doute aussi la plus influente sur les politiques éducatives des pays, même ceux ne participant pas à l'étude ! 28 millions d'élèves dans 70 pays (ou économies) sont évalués, ces pays représentant près de 90% de l'économie mondiale ! 500 000 élèves composent l'échantillon et passent tous les mêmes exercices traduits dans plus de 33 langues !

Le dispositif dans son ensemble est d'une grande qualité scientifique et il y a peu à lui reprocher [8]. Mais comme tout outil de mesure statistique il est soumis à des intervalles de confiance plus ou moins grands. L'intervalle de confiance autour du score global d'un pays dans un domaine PISA dépend de la représentativité de l'échantillon (notamment du nombre d'absents), de la qualité de la correction (mesurée par un système de quadruple correction), et de diverses autres pertes

d'information au cours d'un long processus. Ainsi, dire comme on l'entend parfois [9] que la France se classe en 17^e position sur l'échelle OCDE en culture mathématique n'a pas de sens lorsqu'on considère l'intervalle de confiance autour de son score et celui des autres pays (voir annexe 5). Il est sans doute plus pertinent de comparer les 34 pays de l'OCDE en les classant dans trois groupes, en tenant compte des intervalles de confiance : les pays significativement au-dessus de la moyenne OCDE, les pays sans différence significative avec la moyenne et les pays significativement au-dessous de la moyenne.

BIBLIOGRAPHIE

- [1] Cadre d'évaluation de PISA 2003, OCDE, 2003.
<http://www.oecd.org/dataoecd/46/16/33694924.pdf>
- [2] TIMSS : Trends in International Mathematics and Science Study.
<http://timss.bc.edu/timss2011/index.html>
- [3] CEDRE: Cycle des Évaluations Disciplinaires Réalisées sur Échantillon.
http://media.education.gouv.fr/file/2010/23/9/NIMEN1018_158239.pdf
- [4] ROCHER Thierry, 2008 : Que nous apprennent les évaluations internationales sur le fonctionnement des systèmes éducatifs ? Une illustration avec la question du redoublement.
http://media.education.gouv.fr/file/revue_78/58/1/4_38581.pdf
- [5] Résultats PISA 2009 sur le site de l'OCDE.
http://www.pisa.oecd.org/document/61/0,3746,en_32252351_32235731_46567613_1_1_1_1,00.html
- [6] Note d'information de la DEPP : L'évolution des acquis des élèves de 15 ans en compréhension de l'écrit : premiers résultats de l'évaluation internationale PISA 2009.
http://media.education.gouv.fr/file/2010/99/8/NIMENJVA1024_161998.pdf
- [7] Par exemple NIMIER Jacques : <http://www.pedagopsy.eu/docimologie.html>
- [8] GRENET Julien : http://www.laviedesidees.fr/IMG/pdf/20080208_pisa.pdf
- [9] Par exemple sur le site du Monde :
http://www.lemonde.fr/societe/infographie/2010/12/07/pisa-les-resultats-des-differents-pays_1450385_3224.html

ANNEXE 1 : Item PISA 2003, rendu public en 2006

Cet item contient 170 mots environ : le texte est relativement long. 11% des élèves français de 15 ans (OCDE 9%) ne répondent pas à cette question en 2003. Ce taux de non réponse est élevé pour une QCM et on peut supposer que la quantité de lecture rebute certains élèves qui par conséquent ne tentent pas de répondre.

Par ailleurs, l'amorce, la question, et deux réponses sur quatre (dont la bonne réponse C) sont dans le registre langagier. Et le niveau de langage de la question (« Parmi les propositions suivantes, laquelle exprime le mieux *ce que veut dire ce géologue ?* ») est tellement élevé que les concepteurs ont choisi de mettre en italique la partie de la phrase nécessitant le plus de réflexion de la part de l'élève.

A priori, l'élève doit passer du registre langagier à celui des nombres, du cadre numérique au cadre des probabilités, comparer la fraction « deux sur trois » (et non pas « deux tiers ») avec 0 et 1, connaître la probabilité d'un événement impossible et d'un événement certain et retourner dans le registre langagier pour choisir la réponse C. Environ 51% des élèves français ont choisi cette réponse en 2003 (OCDE 47%).

Question 1 : TREMBLEMENT DE TERRE

On a diffusé un documentaire sur les tremblements de terre et la fréquence à laquelle ils se produisent. Ce reportage comprenait un débat sur la prévisibilité des tremblements de terre.

Un géologue a affirmé : « Au cours des vingt prochaines années, la probabilité qu'un tremblement de terre se produise à Zedville est de deux sur trois. »

Parmi les propositions suivantes, laquelle exprime le mieux *ce que veut dire ce géologue ?*

- A Puisque $\frac{2}{3} \times 20 = 13,3$, il y aura donc un tremblement de terre à Zedville dans 13 à 14 ans à partir de maintenant.
- B $\frac{2}{3}$ est supérieur à $\frac{1}{2}$, on peut donc être certain qu'il y aura un tremblement de terre à Zedville au cours des 20 prochaines années.
- C La probabilité d'avoir un tremblement de terre à Zedville dans les vingt prochaines années est plus forte que la probabilité de ne pas en avoir.
- D On ne peut pas dire ce qui se passera, car personne ne peut être certain du moment où un tremblement de terre se produit.

ANNEXE 2 : Autre item PISA 2003, rendu public en 2006

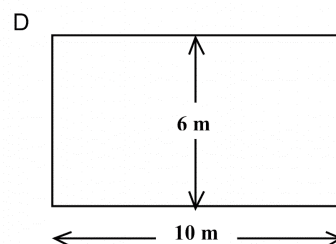
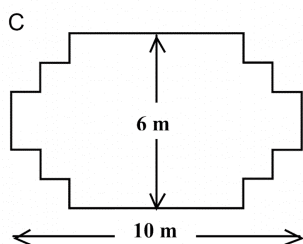
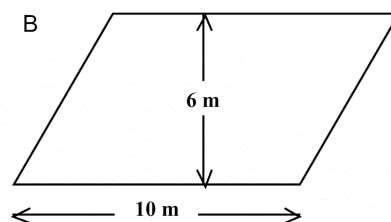
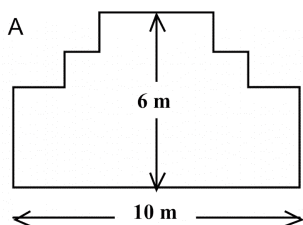
Peu de mots en revanche dans l'énoncé de cet item et pourtant le taux de non réponse des élèves français est 5% (OCDE 2,5%), donc relativement élevé pour un Vrai/Faux. Là encore, des compétences de maîtrise de la langue sont nécessaires : elles ne relèvent pas de la quantité de lecture mais d'interprétation lexicale. A priori, la tâche consiste d'abord à repérer dans le registre langagier de quelle grandeur il est

question : ici seul le mot « bordure » conduit à la notion de périmètre, même si on dispose d'une certaine longueur de planches, donnée sur laquelle l'élève peut inférer. Dans un cadre géométrique, il doit ensuite choisir une stratégie de comparaison des longueurs (mesure, déplacements de segments, superpositions, etc.). Il faut aussi connaître le périmètre du rectangle, le calculer à partir des données et comparer avec 32 dans un cadre numérique.

La tâche est complexe, et le niveau de difficulté élevé : en 2003, 18,5% des élèves français répondent correctement (OCDE 20%).

Question 1 : MENUISIER

Un menuisier dispose de 32 mètres de planches et souhaite s'en servir pour faire la bordure d'une plate-bande dans un jardin. Il envisage d'utiliser un des tracés suivants pour cette bordure :



Indiquez, pour chacun des tracés, s'il peut être réalisé avec les 32 mètres de planches. Répondez en entourant « Oui » ou « Non ».

Tracé de la bordure	En utilisant ce tracé, peut-on réaliser la plate-bande avec 32 mètres de planches ?
Tracé A	Oui / Non
Tracé B	Oui / Non
Tracé C	Oui / Non
Tracé D	Oui / Non

ANNEXE 3 : Les laboratoires cognitifs

Il s'agit de soumettre chaque item en cours de conception à un petit groupe d'élèves (trois idéalement). Les laisser interagir pour résoudre les items et enregistrer leur activité : leurs commentaires oraux, leurs actions ou procédures orales ou écrites de résolution. Les élèves du groupe auront des niveaux divers dans la discipline évaluée.

Plusieurs protocoles sont possibles :

- L'enregistrement vidéo est optimal mais il a la contrainte du droit à l'image requis pour chaque élève.
- L'enregistrement audio est moins contraignant et l'information apportée est susceptible d'être exhaustive.
- Un concepteur est dans tous les cas témoin du travail du groupe et prend note de tout élément pertinent à la conception de l'item. Dans la mesure du possible les élèves sont dans les conditions matérielles de passation et aucune aide pour la résolution (mise à part l'aide de leurs pairs) n'est prodiguée.

Les méthodes de résolution des élèves sont la priorité du concepteur présent afin de vérifier la conformité éventuelle avec ce que le panel avait prévu. Comme il assiste en direct à la résolution il est témoin des démarches et raisonnements mis en place par les élèves et en prend note.

Les écueils rencontrés par les élèves en termes de formulation, de précision des données (par exemple les unités utilisées), de degré d'explicitation des amorces et questions, etc. seront aussi notés.

Le temps passé sera enregistré au cours de la résolution.

En fin de passation, alors que la résolution est terminée ou n'a pas pu aboutir, un questionnaire spécifique est soumis aux élèves. Ce questionnaire a posteriori fait le point avec eux sur leurs réactions générales.

Le panel de concepteurs se réunit à nouveau pour analyser les résultats des laboratoires cognitifs. Les notes du concepteur, les enregistrements audio ou vidéo sont analysés. Les modifications sont apportées aux items en conséquence. Elles pourront notamment porter sur : l'estimation de la durée de l'item, de la difficulté de l'item, de la motivation de l'élève face à l'item (notion d'intérêt, d'ennui), sur les distracteurs⁽²⁾ de QCM, la pertinence du format choisi pour l'item (QCM, Vrai/Faux, ouvert, ...), la lisibilité du texte et des informations, le choix d'un item plutôt qu'un autre pour une même situation, les consignes de correction/codage.

ANNEXE 4 : Consignes de codage/correction

La conception d'un item ne peut se faire sans considération des consignes de correction ou de codage. Ces consignes déterminent ce que l'item évalue. Elles allouent certains « crédits⁽³⁾ » à certaines réponses d'élève, aucun crédit à d'autres.

(2) Dans une question à choix multiple, on appelle distracteur toute réponse proposée (et donc présentant une certaine vraisemblance) en dehors de la bonne réponse.

(3) Le « crédit », terme international d'origine américaine, correspond ici à un code numérique.

Ces crédits sont de trois sortes : crédit complet, partiel ou aucun crédit.

Le crédit complet est alloué à la réponse attendue dans le cas où l'élève a résolu l'item au mieux de ce qu'il évalue. Dans le cas où le format de l'item est de type ouvert, la description de l'attendu pour un crédit complet peut être très détaillée tenant compte du fait que les correcteurs, bien qu'experts, ne sont pas toujours des concepteurs d'item et que ce qui peut sembler sous-entendu pour l'un ne l'est pas forcément pour l'autre !

Un crédit complet peut être multiple. Par exemple, certains items peuvent admettre plusieurs méthodes ou pistes de résolution équivalentes et un crédit complet sera par conséquent alloué à chacune de ces réponses, de manière exhaustive. Ce cas se rencontre surtout dans un item de format ouvert. Le crédit complet dépend de la compétence ou connaissance pour l'évaluation desquelles l'item a été construit par les concepteurs. L'élève peut donc obtenir un crédit complet alors qu'une partie de sa réponse comporte une erreur. Le concepteur peut décider de ne pas tenir compte de l'erreur si elle ne dépend pas de ce qui est évalué dans l'item et pourra le préciser en développant une consigne de crédit complet multiple.

Un crédit partiel sera alloué à une réponse ne validant que partiellement ce que l'item évalue et notamment dans les cas suivants. Lorsque plusieurs connaissances ou compétences ou étapes de raisonnement sont en jeu et que la réponse montre qu'une partie seulement a été mobilisée par l'élève. Lorsqu'une procédure utilisée par l'élève est considérée comme moins experte que celle attendue par les concepteurs. Lorsqu'une erreur a pu être commise à un stade du raisonnement mais que la réponse montre que les autres stades du raisonnement ont été effectués correctement.

La description des crédits partiels d'un item dépend des compétences ou connaissances que le concepteur désire évaluer avec l'item. En les rédigeant, le concepteur doit toujours avoir en ligne de mire l'analyse a posteriori des résultats du test. Un crédit partiel doit donc permettre une analyse plus fine des acquis des élèves sans pour autant la compliquer inutilement.

Il est à noter ici que le développement de consignes de codage et notamment les descriptions de l'attendu d'un crédit complet ou partiel n'est jamais définitif avant la dernière phase de conception. Un laboratoire cognitif par exemple peut apporter une modification aux crédits partiels, les élèves donnant une réponse pertinente ou envisageant une piste de résolution que les concepteurs n'avaient pas prévues. Ce sont surtout les tests à 50 élèves qui détermineront la pertinence ou pas de certains crédits partiels.

L'exemple suivant, CONTINENT, est tiré du test PISA 2000 en culture mathématique. Le détail des consignes de correction montre la forte volonté de standardisation en rapport avec une analyse fine de l'activité des élèves. Les consignes sont présentées ici telles que les correcteurs PISA du monde entier (modulo la langue utilisée) peuvent les lire.

Question 2 : CONTINENT

Estimez l'aire de l'Antarctique en utilisant l'échelle de cette carte.

Montrez votre travail et expliquez comment vous avez fait votre estimation. (Vous pouvez dessiner sur la carte si cela vous aide pour votre estimation.)

**Consignes de correction****Crédit complet**

[Les codes suivants sont à attribuer aux réponses où l'approche utilisée ET le résultat sont corrects. Le second chiffre du code sert à distinguer les différentes approches]

- Code 21 : Estime l'aire en dessinant un carré ou un rectangle – réponse comprise entre 12 000 000 et 18 000 000 km carrés (les unités ne sont pas requises).
- Code 22 : Estime l'aire en dessinant un cercle – réponse comprise entre 12 000 000 et 18 000 000 km carrés.
- Code 23 : Estime l'aire en additionnant l'aire de plusieurs figures géométriques régulières – réponse comprise entre 12 000 000 et 18 000 000 km carrés.
- Code 24 : Estime l'aire de manière correcte en utilisant une autre méthode – réponse comprise entre 12 000 000 et 18 000 000 km carrés.
- Code 25 : Réponse correcte (comprise entre 12 000 000 et 18 000 000 km carrés), mais pas d'indication sur la méthode utilisée.

Crédit partiel

[Les codes suivants sont à attribuer aux réponses où l'approche utilisée est correcte, mais le résultat est incorrect ou incomplet. Le second chiffre du code sert à distinguer les différentes approches, et correspond au second chiffre du code utilisé pour le crédit complet]

- Code 11 : Estime l'aire en dessinant un carré ou un rectangle – méthode correcte, mais réponse incomplète ou erronée.
Dessine un rectangle et multiplie la largeur par la longueur – méthode correcte, mais la réponse surestime ou sous-estime l'aire (par exemple : 18 200 000).
Dessine un rectangle et multiplie la largeur par la longueur, mais le nombre de zéros est incorrect (par ex., $4\,000 \times 3\,500 = 140\,000$).
Dessine un rectangle et multiplie la largeur par la longueur, mais oublie d'utiliser l'échelle pour convertir le résultat en km carrés (par ex., $12\text{ cm} \times 15\text{ cm} = 180$).
Dessine un rectangle et indique que l'aire est de $4\,000\text{ km} \times 3\,500\text{ km}$, sans présenter la suite du travail.
- Code 12 : Estime l'aire en dessinant un cercle – méthode correcte, mais réponse incomplète ou erronée.
- Code 13 : Estime l'aire en additionnant l'aire de plusieurs figures géométriques régulières – méthode correcte, mais réponse incomplète ou erronée.
- Code 14 : Estime l'aire en utilisant une autre méthode correcte – mais réponse incomplète ou erronée.

Pas de crédit

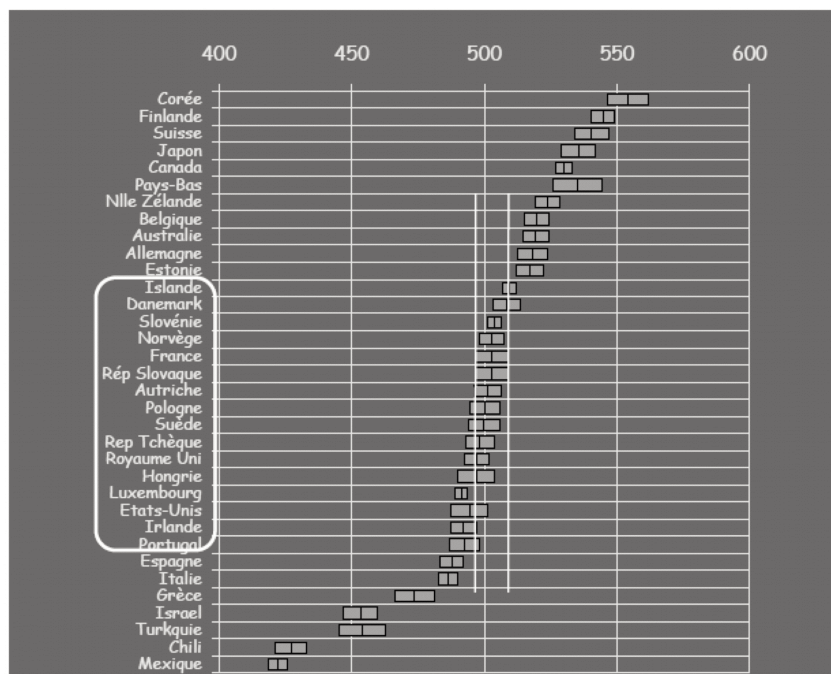
- Code 01 : Calcule le périmètre au lieu de l'aire.
Par ex., 16 000 km, parce que l'échelle de 1 000 km va 16 fois autour de la carte.

Code 02 : Autres réponses incorrectes.

Par ex., 16 000 km [le raisonnement n'est pas montré, et la réponse est incorrecte].

Code 99 : Omission.

ANNEXE 5 : Scores globaux (et intervalle de confiance) des pays de l'OCDE au PISA 2009 en culture mathématique.



Lecture : Compte tenu de l'intervalle de confiance autour du score de la France (lignes blanches) et des intervalles de confiance autour des scores des autres pays de l'OCDE, le score de la France en culture mathématique en 2009 n'est pas significativement différent des scores des pays encadrés en blanc (hormis le Luxembourg).

Source OCDE 2009