



6

Cadre conceptuel des questionnaires contextuels du cycle PISA 2012

Le cadre conceptuel des questionnaires contextuels du cycle PISA 2012 montre que les résultats PISA peuvent constituer une base de données durable qui soit pertinente pour la politique de l'éducation et la recherche en pédagogie. Il décrit les fondements à retenir pour que les questionnaires contextuels PISA restent probants à l'avenir. Il formule des recommandations à propos de la structure et de l'analyse pour exploiter au mieux les points forts de l'enquête PISA et jeter des bases encore plus solides pour l'avenir. Enfin, il expose les fondements théoriques d'une grande rigueur scientifique des concepts, des échelles et des indices du cycle PISA 2012 qui seront élaborés à partir des différents questionnaires contextuels administrés dans les pays et les économies participants. Ce cadre passe en revue de manière détaillée de nouveaux indicateurs de résultat individuel – les stratégies, les convictions et la motivation – en rapport avec la culture mathématique, ainsi que des indicateurs innovants sur les possibilités d'apprentissage et la qualité de l'instruction au niveau des élèves, des établissements d'enseignement et des systèmes d'éducation.



INTRODUCTION

Depuis sa création à la fin des années 90, le Programme international de l'OCDE pour le suivi des acquis des élèves (PISA) est connu pour sa contribution significative aux débats sur la politique de l'éducation des pays de l'OCDE et des pays et économies partenaires. La « littératie » et les compétences utiles dans la vie, qui sont évaluées à l'âge de 15 ans, reflètent le rendement cumulé de l'apprentissage dans le cadre scolaire et en dehors. La compréhension de l'écrit, la culture mathématique et la culture scientifique sont des atouts importants pour la réussite académique, scolaire, civique, sociale et privée des individus. Rapporté par système ou par pays, le niveau moyen de ces compétences est un indicateur probant du capital humain, déterminant pour la prospérité et le bien-être de la société. La répartition de ces compétences dans la population en fonction du sexe, de l'origine, du milieu socio-économique ou de la situation géographique contribue à expliquer la diversité et l'équité entre les pays et au sein même de ceux-ci. Les résultats aux épreuves cognitives PISA sont combinés à des informations sur la situation des élèves pour fournir des informations sur la diversité, l'équité et le niveau des compétences utiles dans la vie. Les décideurs, les chercheurs, les professionnels de l'éducation, les parents, les élèves et l'opinion au sens large peuvent comparer les résultats de leur pays à ceux d'autres pays en utilisant les données PISA comme valeurs de référence. Pour recueillir les informations contextuelles requises, les élèves sélectionnés lors de tous les cycles de l'enquête PISA sont questionnés à divers égards, notamment leur sexe, leur origine (autochtone ou allochtone), et le niveau de formation et la profession de leurs parents.

La littératie et les compétences utiles dans la vie sont en grande partie le fruit de l'éducation dans le milieu familial et le cadre scolaire. Pour décrire ces contextes d'apprentissage et mieux cerner leur impact sur la performance des élèves, l'enquête PISA demande aussi aux élèves, aux chefs d'établissement et, parfois, aux parents d'élèves de faire part de leur point de vue sur ces contextes. Ces informations supplémentaires nous aident à répondre à diverses questions. La culture pédagogique varie-t-elle entre les pays, soit, en d'autres termes, pouvons-nous identifier des profils pédagogiques spécifiques aux pays ? Des établissements d'enseignement différents accueillent-ils à l'échelle nationale des effectifs d'élèves différents, soit, en d'autres termes, leurs politiques et pratiques ciblent-elles leurs effectifs respectifs ? Les écarts de performance entre les élèves peuvent-ils s'expliquer par des différences dans leurs caractéristiques sociétales, leur famille, les ressources et les pratiques pédagogiques de leur établissement, et leur communauté ? Les réponses apportées à ces questions par l'enquête PISA ont fait grand bruit dans les débats sur la qualité, l'efficacité et l'efficacité des systèmes d'éducation dans le monde. En une décennie, les cycles PISA 2000, 2003, 2006 et 2009 ont permis de créer l'une des bases de données les plus importantes du monde qui peut être utilisée afin de tirer des enseignements pour éclairer l'action publique et d'évaluer l'efficacité de l'instruction. Les responsables politiques peuvent se servir des réponses qu'ils trouvent dans les résultats de l'enquête PISA pour réorienter leur action, redéfinir leurs objectifs, revoir les priorités de la réforme de l'enseignement, réaffecter les ressources, promouvoir de nouveaux concepts pédagogiques et de nouvelles stratégies organisationnelles et, en fin de compte, améliorer l'éducation. Comme l'enquête PISA réévalue les systèmes d'éducation tous les trois ans, elle éclaire les changements dans l'éducation et en suit l'évolution.

Comme l'enquête PISA en est à son cinquième cycle, le moment est venu de revoir le cadre conceptuel qui oriente la conception des questionnaires et la façon dont les informations sont utilisées pour analyser et rendre compte des données. Entre 2000 et 2009, par exemple, le questionnaire « Élève » a permis de recueillir des données sur plus de 70 indicateurs (échelles agrégées à partir des données du questionnaire) à propos des activités et des ressources familiales, du cadre d'apprentissage en milieu scolaire et à domicile, et des perceptions et des motivations des élèves. Toutefois, aucun de ces indicateurs n'a été administré lors des quatre cycles, et la majorité d'entre eux (plus de 55) ne l'ont été qu'une seule fois. Les cycles PISA ont donc jusqu'à un certain point été conçus comme une étude indépendante sur les contextes d'apprentissage axée sur un domaine majeur spécifique. Le contenu des questionnaires a été largement déterminé par des priorités ponctuelles plutôt que par des objectifs à long terme en matière de recherche et d'action publique. Ce principe de conception s'est incontestablement révélé utile pendant que l'enquête était en cours de développement et que de nouveaux domaines ou indicateurs continuaient d'apparaître.

Le cadre conceptuel et analytique de l'enquête continuera assurément d'évoluer en fonction des changements dans la société et dans l'éducation – par exemple, l'importance croissante des technologies de l'information et de la communication (TIC) en tant qu'instrument d'apprentissage et domaine de compétences transversales, ou la nécessité de combiner plusieurs cadres d'apprentissage (à l'école, pendant des activités extrascolaires, par e-learning et même l'éducation à domicile). L'enquête PISA en est toutefois arrivée à un certain degré de maturité. L'enquête PISA étant une source probante et stable de comparaisons internationales et d'informations tendanciennes sur les systèmes d'éducation, la mise au point des prochains questionnaires requiert l'adoption d'une structure cohérente, qui concilie stabilité et innovation, tout en combinant thématiques générales et thématiques spécifiques aux domaines d'évaluation. Ce document décrit la structure qui a été utilisée pour concevoir les questionnaires contextuels du cycle PISA 2012.



Ce chapitre se divise en cinq grandes sections :

- La première section (« L'objectif général de l'enquête PISA et sa pertinence pour l'action publique ») décrit l'objectif général de l'enquête et ses objectifs spécifiques à l'action publique, et explique les types d'informations contextuelles nécessaires pour réaliser ses objectifs et les raisons de leur utilité. L'enquête PISA doit s'intéresser à divers facteurs en rapport avec les élèves et leur famille, avec les établissements d'enseignement et avec le système d'éducation puisqu'elle procède à un suivi comparatif, cherche à comprendre l'efficacité de la politique de l'éducation et des pratiques pédagogiques, et alimente les recherches sur l'action publique. Les questions relatives à l'égalité des chances appellent une attention toute particulière dans la politique de l'éducation.
- La deuxième section (« La base de connaissances générales : la recherche sur l'efficacité de l'éducation ») décrit la base de connaissances générales que l'enquête PISA peut utiliser pour remplir ses missions. Elle commence par présenter la structure qui permet d'organiser les *constructs* et les indicateurs à différents niveaux. La taxinomie s'inspire de la littérature sur l'efficacité de l'éducation, qui est décrite de façon détaillée pour justifier la sélection des intrants, processus et extrants de niveau Système, Établissement et Élève. Des conclusions d'ordre général sur l'efficacité et la productivité de l'éducation y sont également présentées.
- La troisième section (« Les conditions d'apprentissage au service de la culture mathématique ») étudie les environnements d'apprentissage en rapport avec les mathématiques, domaine majeur d'évaluation du cycle PISA 2012. Les questionnaires sont conçus pour permettre l'inclusion d'autres indicateurs de résultats, par exemple les attitudes et les stratégies, la qualité de l'instruction en mathématiques – y compris la possibilité que les élèves aient été exposés aux types de problèmes retenus dans les épreuves PISA (notamment les possibilités d'apprentissage) – et l'ajout de variables de niveau Système et Établissement.
- La quatrième section (« Conception des questionnaires du cycle PISA 2012 ») décrit les questionnaires administrés lors du cycle PISA 2012, qui reposent sur cette base de connaissance et dont le contenu est déterminé par les objectifs de l'enquête PISA dans le domaine de l'action publique. En premier lieu, elle présente la structure globale des *constructs* et indicateurs qui sera appliquée lors des prochains cycles PISA. En deuxième lieu, elle revisite les indicateurs utilisés lors du cycle PISA 2003, le premier à avoir les mathématiques comme domaine majeur d'évaluation, dont la pertinence et la qualité sont établies pour tirer parti des travaux antérieurs et décrire des tendances. Enfin, elle évoque des questions de mise en œuvre inédites dans l'enquête PISA, en l'occurrence au sujet de l'adoption de la rotation des items dans le questionnaire « Élève », qui permettra d'étoffer l'instrument. Certaines questions sont soumises à tous les élèves, comme lors des cycles précédents, et d'autres le sont uniquement à des sous-groupes d'élèves. Quant au questionnaire « Établissement », son administration informatisée améliorera la convivialité de l'instrument et son taux de réponse.
- La cinquième section (« Garantir la validité des instruments PISA, leur pouvoir explicatif et leur pertinence pour l'action publique ») étudie des principes fondamentaux de la structure et de l'analyse qui pourraient contribuer à préserver le caractère novateur de l'enquête PISA et la validité de ses mesures à l'avenir : les techniques de modélisation statistique utilisées pour définir et identifier des effets, garantir l'équivalence et la validité des mesures entre les cultures, l'échantillonnage d'établissements d'enseignement déjà échantillonnés lors du cycle PISA 2003 pour suivre l'évolution à l'échelle des établissements depuis le cycle PISA 2003 et l'utilisation d'études de suivi ou d'extensions longitudinales pour quantifier la valeur ajoutée de l'éducation.

Les questionnaires communs et les questionnaires proposés à titre d'option internationale lors du cycle PISA 2012 sont repris à l'annexe A.

L'OBJECTIF GÉNÉRAL DE L'ENQUÊTE PISA ET SA PERTINENCE POUR L'ACTION PUBLIQUE

Objectifs majeurs de l'enquête PISA

L'enquête PISA poursuit un éventail d'objectifs parallèles. Les points de vue des différentes parties prenantes dans les pays et économies participants se répartissent entre les domaines suivants :

- L'enquête PISA est un dispositif de suivi qui fournit des informations comparatives fiables sur les systèmes d'éducation, dont elle décrit les structures, le fonctionnement et la productivité (c'est-à-dire leur « rendement »). Les données PISA couvrent la performance des élèves, leur motivation et leur parcours scolaire jusque dans l'enseignement secondaire, les caractéristiques des établissements, la gouvernance scolaire et les questions en rapport avec l'équité (par exemple, la variation de la performance selon le sexe ou le milieu socio-économique).
- L'enquête PISA est une étude internationale qui contribue à enrichir notre base de connaissances sur l'efficacité de l'éducation. Elle permet d'observer les modèles de relations entre les intrants, les processus et les résultats de



l'éducation. Elle offre la possibilité de décomposer la variation de la performance des élèves et de chiffrer la part de cette variation qui est imputable à des facteurs spécifiques aux élèves, aux établissements d'enseignement et au système d'éducation. Chaque cycle de l'enquête PISA fournit de surcroît des données sur de nombreux facteurs – en rapport avec ces trois niveaux – censés, selon des recherches antérieures, influencer sur la performance en compréhension de l'écrit, en mathématiques et en sciences. L'enquête PISA décrit ces facteurs et estime leur impact direct ou indirect sur la performance des élèves et d'autres résultats. Le prélèvement de grands échantillons représentatifs et l'utilisation de modèles statistiques sophistiqués permettent de généraliser les résultats tant dans les pays qu'entre eux.

- L'enquête PISA est une source de données qui permet d'étudier les contextes d'apprentissage en général (par exemple, les interactions de l'éducation dans le milieu familial, scolaire et extrascolaire, et leur impact sur le développement des compétences utiles dans la vie), et les variables en rapport avec l'éducation dans des contextes économiques et sociologiques (par exemple, la relation entre la démographie, la richesse, la croissance économique et les ressources humaines). La base de connaissances est étoffée par des données tendanciennes sur les mathématiques, et couvrira 12 années dès que les résultats du cycle PISA 2012 y seront inclus.

L'enquête PISA propose trois types de « produits » pertinents pour l'action publique : *i*) des indicateurs de suivi du fonctionnement, de la productivité et de l'équité des systèmes d'éducation ; *ii*) des informations sur les facteurs déterminants de l'efficacité de l'éducation ; et *iii*) une base de données fiable et durable se prêtant aux comparaisons qui permet aux chercheurs du monde entier d'approfondir leurs études sur des questions de pédagogie ainsi que sur des questions en rapport avec l'action publique.

Suivi comparatif des systèmes

L'un des objectifs majeurs de l'enquête PISA est de suivre les systèmes d'éducation à plusieurs égards : la performance des élèves (en termes de littératie et, plus généralement, de compétences utiles dans la vie), certains de leurs résultats non cognitifs (leur motivation et leur bien-être), leur parcours scolaire et certaines de leurs caractéristiques contextuelles (leur appartenance ethnique et leur milieu culturel et socio-économique), ainsi que des caractéristiques propres aux établissements d'enseignement et au système d'éducation (notamment les politiques d'évaluation et de responsabilisation, la sélection et l'affectation des élèves, l'implication des parents, la coopération entre les membres du personnel et les possibilités d'apprentissage). L'enquête PISA fournit également des indicateurs qui révèlent les relations entre ces facteurs, notamment le gradient social qui mesure l'intensité de la relation entre le milieu socio-économique et la performance, ou la relation entre les intrants et les extrants de l'éducation. Parmi les usages faits des données PISA, le plus important consiste à élaborer des indicateurs sur les intrants, les processus et les extrants, ainsi que des indicateurs relationnels, qui alimentent le programme de l'OCDE sur les indicateurs de l'éducation qui, à son tour, alimente le débat public, façonne la politique de l'éducation et éclaire la prise de décision.

L'enquête PISA s'inscrit dans une perspective internationale. Elle est conçue pour doubler les efforts consentis à l'échelle nationale et régionale en vue de recueillir des informations détaillées sur les élèves, les établissements d'enseignement et les systèmes d'éducation. Elle complète ces sources en fournissant des valeurs de référence internationales et des exemples de ce qui se fait dans le monde entier. À l'heure de la mondialisation économique et sociale, ces exemples peuvent donner de nouvelles idées et de nouveaux éléments pour orienter la politique de l'éducation à l'échelle nationale, régionale ou locale.

Ce système de suivi doit sa pertinence par rapport à l'action publique à plusieurs principes : *i*) la définition et l'opérationnalisation d'indicateurs de résultats cognitifs et non cognitifs qui permettent de choisir et de hiérarchiser les priorités de l'éducation dans les pays participants ; *ii*) l'analyse et la présentation de facteurs sur lesquels l'action publique et les pratiques pédagogiques peuvent influencer (les facteurs dits « malléables ») ; et *iii*) l'élaboration de valeurs de référence internationales, qui permettent aux responsables politiques de tirer des enseignements de ce que font d'autres pays. Les indicateurs sont en général sélectionnés à la lumière des demandes des responsables politiques. Les responsables de la politique de l'éducation doivent s'attacher au fonctionnement du système d'éducation (notamment à des caractéristiques telles que les ressources allouées aux établissements d'enseignement), à sa productivité (par exemple, les résultats bruts des élèves) et, aspect important s'il en est, à son équité (dont la répartition des ressources).

Pour présenter des indicateurs pertinents pour l'action publique, il faut non seulement évaluer la performance des élèves, en mathématiques par exemple, mais aussi analyser les données recueillies via les questionnaires « Élève », « Établissement » et « Parents » qui couvrent un large éventail d'attributs en rapport avec des contextes, des processus et des résultats non cognitifs. Ce document présente les arguments théoriques et les considérations propres à l'action publique dont il est tenu compte pour sélectionner les *constructs* et définir les questions et les échelles.



Avant l'enquête PISA, un certain nombre d'études quantitatives et qualitatives sur l'éducation ont permis de recueillir des éléments sur l'historique, le fonctionnement et, jusqu'à un certain point, sur l'efficacité des systèmes d'éducation. L'enquête PISA est toutefois unique en son genre, car elle combine les atouts suivants :

- Elle fournit des données rigoureuses et une série intégrée d'indicateurs qui permettent de suivre les systèmes d'éducation.
- Elle se concentre à dessein sur le rendement, c'est-à-dire la performance des élèves à l'approche de la fin de la scolarité obligatoire. Elle cherche également à décrire des attributs malléables des systèmes d'éducation susceptibles de réorienter l'action publique, la mise en œuvre d'interventions et l'introduction d'améliorations.
- Elle fournit des indicateurs tendanciels, qui permettent de décrire l'évolution du niveau de performance des pays au fil du temps, élabore des variables sur des contextes, des processus et des résultats non cognitifs, et examine les relations entre ces divers *constructs*. Plus les cycles PISA se succèdent, plus des enseignements peuvent être tirés de l'examen de la stabilité ou de l'évolution des intrants, des processus et des extrants de l'éducation, et de la variation de leurs relations dans le temps.

De nombreux exemples d'indicateurs dérivés des données contextuelles de l'enquête PISA se trouvent dans les éditions récentes du rapport de l'OCDE *Regards sur l'éducation* (OCDE, 2007a, 2008, 2009a, 2011, 2012a), notamment :

- Les relations entre l'ascendance allochtone ou le statut socio-économique, d'un côté, et la performance des élèves, de l'autre (2007, indicateur A6 ; 2011 et 2012, indicateur A5).
- Le profil des élèves les plus performants en sciences, notamment leurs attitudes et leurs motivations à l'égard des sciences (2009, indicateurs A4 et A5).
- Les relations entre le plaisir de la lecture et la performance en compréhension de l'écrit (2007, indicateur A5).
- Les relations entre les ressources et les résultats de l'éducation (2007 et 2008, indicateur B7), en particulier compte tenu de la taille des classes (2008, indicateur D2).
- La comparaison du rendement de l'éducation entre la filière générale et la filière professionnelle (2007 et 2008, indicateur C1).
- L'utilisation des évaluations et des examens dans les systèmes d'éducation (2008, indicateur D5).
- La qualité des établissements selon les parents (2008, indicateur A6),
- Les relations entre le milieu des élèves et leur accès à l'enseignement supérieur (ou leur motivation en la matière) (2007, indicateurs A4 et A7 ; 2008, indicateurs A3 et A7).

L'un des défis les plus importants à relever concernant l'élaboration du cadre conceptuel des questionnaires contextuels consiste à garantir que les indicateurs sont comparables entre les cycles, tout en continuant à introduire de nouveaux indicateurs. Après une décennie de cycles PISA et d'évaluation internationale de la performance des élèves, le moment est venu de structurer et d'ordonner tous les *constructs* et instruments qui peuvent être utilisés soit ponctuellement, soit durablement à titre de source de données tendancielles. La pertinence de l'enquête PISA à l'avenir en dépend assurément.

Comprendre les modèles d'efficience dans les systèmes d'éducation

Des indicateurs attireront l'attention de l'opinion sur des aspects fructueux ou moins fructueux du système d'éducation, ou sur des objectifs qui ont été atteints ou qui ne l'ont pas encore été. L'objectif principal est donc d'orienter la hiérarchisation des priorités et la prise de décisions. Les responsables politiques veulent non seulement qu'on leur explique les points forts et les points faibles des élèves, et les conditions d'enseignement et de scolarisation dans leur pays, mais également comprendre pourquoi les élèves atteignent tel ou tel niveau de compétence. Les instruments contextuels PISA sont conçus pour contribuer à répondre à leurs questions. Les questionnaires PISA doivent donc couvrir les intrants et les processus les plus importants de l'apprentissage des élèves au niveau individuel, ainsi qu'à l'échelle des établissements d'enseignement et du système d'éducation. Les modèles statistiques contribuent, sur la base de ces données multiniveau, à mieux comprendre la complexité des interactions entre ces intrants et processus et les résultats des élèves. Lorsque les coûts et ressources sont connus, l'enquête PISA aide également à évaluer l'efficience, c'est-à-dire le retour sur investissement.

Durant la première décennie de l'enquête PISA, chaque cycle s'est concentré sur des variables en rapport avec le domaine majeur d'évaluation. En compréhension de l'écrit, par exemple, les expériences de lecture, l'intérêt pour la lecture, l'enseignement en classe, etc. ont été privilégiés en 2000 et en 2009. Le cycle PISA 2003 a accordé la priorité aux attitudes à l'égard des mathématiques et aux activités en cours de mathématiques, alors que le cycle PISA 2006



a exploré les diverses pédagogies en sciences et les attitudes et les motivations des élèves en la matière. L'*indice du climat de discipline en classe* et l'*indice du soutien apporté par les enseignants en classe* ont été mesurés en langue de l'évaluation en 2000 et en 2009, et en mathématiques en 2003. Les données ont été opérationnalisées et analysées en fonction des domaines d'évaluation, mais les *constructs*, les relations théoriques avec les résultats des élèves et l'interprétation des données ont été définis compte tenu de théories générales sur l'efficacité de l'éducation.

Ces données propres à chaque domaine d'évaluation ont permis de mieux cerner certains aspects, très importants pour la politique de l'éducation et les pratiques pédagogiques. L'OCDE clôture chaque cycle PISA par la rédaction d'un rapport initial (OCDE, 2001 ; OCDE, 2004 ; OCDE, 2007 ; OCDE, 2010). L'OCDE publie également à l'occasion de chaque cycle une série de rapports thématiques, c'est-à-dire des analyses approfondies des données PISA sur un large éventail de sujets concernant les élèves, les établissements d'enseignement et le système d'éducation (parmi les rapports récents, citons *Untapped Skills: Realising the Potential of Immigrant Students*, OCDE, 2012b ; *Public and Private Schools: How Management and Funding Relate to their Socio-economic Profile*, OCDE, 2012c ; *Lisons-leur une histoire ! Le facteur parental dans l'éducation*, OCDE, 2012d). Depuis 2012, l'OCDE publie également *PISA à la loupe*, une série de notes de synthèse mensuelles sur la politique de l'éducation qui décrivent brièvement un sujet pertinent pour l'action publique¹. L'OCDE vient aussi de lancer la série *Strong Performers and Successful Reformers in Education* pour prodiguer des conseils taillés sur mesure aux pays qui s'adressent à l'OCDE ou pour montrer la réussite du système d'éducation très performant d'un pays afin de permettre à d'autres pays de s'en inspirer².

Au-delà des publications de l'OCDE, voici quelques exemples intéressants tirés de la première conférence PISA qui a eu lieu en septembre 2009³. Certains exemples sont basés sur des analyses internationales, alors que d'autres reposent sur les données PISA d'un seul pays :

- L'élévation du niveau d'engagement des élèves à l'égard de la lecture permet d'améliorer l'équité (Baye *et al.*, 2009).
- Les moyens affectés aux établissements d'enseignement et la mesure dans laquelle ces derniers encouragent les élèves à exploiter tout leur potentiel sont des variables prédictives probantes de la performance en mathématiques en Corée (Kaplan, 2009a).
- Le temps d'instruction est en forte corrélation avec la performance en mathématiques dans les établissements d'enseignement suisses, sauf dans ceux où les exigences sont élevées (Angelone *et al.*, 2009).
- Pour favoriser le développement des compétences des élèves en sciences et leur intérêt pour la science, polariser les cours sur des thématiques scientifiques semble être plus judicieux qu'élargir la gamme des activités scientifiques (Kobarg *et al.*, 2009).
- La mesure dans laquelle les établissements d'enseignement prédisposent à l'exercice d'une profession scientifique est en forte corrélation avec la volonté des élèves d'exercer une telle profession (Lie et Kjaernsli, 2009).
- Les élèves japonais s'intéressent à l'apprentissage inquisitif, mais les professeurs de sciences ne tiennent pas compte de cet intérêt dans l'enseignement secondaire. Ce constat pourrait aider à expliquer pourquoi les élèves japonais accusent des faibles niveaux d'attitudes positives à l'égard de la science (Yasushi, 2009).

L'analyse des données PISA peut largement contribuer à enrichir la base de connaissances sur la politique de l'éducation et les pratiques pédagogiques, mais certaines limites doivent être prises en considération. La plus importante réside dans le fait que l'enquête PISA est une étude de rendement qui évalue les connaissances et compétences cumulées depuis la petite enfance jusqu'à l'âge de 15 ans, au fil des différents niveaux d'enseignement. L'enquête PISA ne cherche pas à déterminer ce que les élèves ont appris dans l'établissement d'enseignement qu'ils fréquentent au moment de l'administration des épreuves. Pour ce faire, il faudrait commencer par évaluer la performance des élèves au moment de leur entrée dans cet établissement et la comparer avec leur performance actuelle. Cela permettrait d'évaluer la progression de la performance associée à l'expérience d'apprentissage vécue dans cet établissement. Or, nous ne disposons pas d'indicateurs sur la performance des élèves avant l'âge de 15 ans, l'âge cible de l'enquête PISA. L'enquête PISA ne permet pas non plus – du moins en fonction de la structure en usage depuis plus de 10 ans – de juger de la qualité des enseignants et de son impact sur la performance des élèves. En effet, elle est administrée à des échantillons de sujets prélevés de façon aléatoire parmi les élèves âgés de 15 ans dans chaque établissement d'enseignement et non à des classes entières, ce qui empêche de recueillir des données sur les stratégies pédagogiques à l'échelle des classes. Enfin, les élèves viennent de commencer à fréquenter leur établissement d'enseignement dans un pays sur cinq parmi ceux qui ont participé au cycle PISA 2006 : il est donc pratiquement impossible de tirer des conclusions directes sur l'effet des établissements d'enseignement dans ces pays.



Comme Baker (2009) le souligne, l'historique de l'élaboration des politiques à la lumière d'études comparatives internationales montre que des conclusions hâtives ont parfois été tirées à cause d'hypothèses trop simples sur les causes à l'origine des écarts de performance à l'échelle des systèmes. Par ailleurs, des économétriciens ont étudié un certain nombre d'aspects de la productivité de l'éducation, mais l'essentiel de leurs travaux est de nature descriptive et ne permet pas de faire des inférences causales (voir Hanushek et Woessmann, 2010).

Il est extrêmement difficile de dégager des inférences causales, par exemple conclure que telle ou telle politique d'éducation ou pratique pédagogique a un impact direct ou indirect sur la performance des élèves, de données issues d'observations et d'évaluations comme celles proposées dans l'enquête PISA (Gustafsson, 2007 ; Kaplan, 2009b). Si des liens étaient établis entre la performance élevée des élèves et les données publiques sur les évaluations des établissements d'enseignement (dans le cadre de la politique sur les établissements, par exemple), la structure de l'enquête ne permettrait pas de les interpréter sous forme de relations causales. Cela s'explique par le fait que certains facteurs potentiellement importants, tels que la performance antérieure des élèves, ne peuvent être recueillis dans l'enquête PISA. En conséquence, il est impossible d'inclure ces facteurs potentiellement importants dans les analyses, ni de les contrôler de façon statistique. L'enquête PISA ne permet donc pas d'affirmer si cette politique n'est appliquée que dans les établissements d'enseignement très performants ou si elle entraîne effectivement une augmentation de la performance des élèves. Le problème fondamental réside dans le fait qu'en l'absence d'affectation aléatoire à un traitement (une politique ou une pratique), il n'est en général pas possible de déterminer si des facteurs non observés influent sur la sélection du traitement ou sur le résultat à l'étude. La possibilité de faire des inférences causales dépend donc de la volonté des chercheurs de formuler des hypothèses supplémentaires, qu'il leur est souvent impossible de tester. Certains chercheurs peuvent essayer de comparer des variables de contrôle observées entre les élèves témoins et les élèves traités pour écarter les élèves qui ne s'y conforment pas, mais le problème reste entier : il est possible que la sélection soit guidée par des variables non observées. En tout état de cause, il est essentiel que les chercheurs formulent clairement toutes les hypothèses qu'ils ont faites et, idéalement, qu'ils évaluent dans quelle mesure leurs inférences causales sont sensibles aux biais de ces hypothèses. (Voir la section « Garantir la validité des instruments PISA, leur pouvoir explicatif et leur pertinence pour l'action publique » pour un exposé analytique détaillé des méthodes et structures visant à remédier à certains de ces problèmes.)

La valeur du programme repose en grande partie sur les interactions constantes entre l'étude de suivi qu'est l'enquête PISA et des recherches plus rigoureuses sur l'efficacité effectuées par ailleurs. Les résultats corrélacionnels et exploratoires de l'enquête PISA peuvent ensuite être mis à l'épreuve dans des études longitudinales ou expérimentales, ou dans des analyses d'interventions. Des facteurs déclarés pertinents pour l'efficacité de l'éducation dans la littérature sont très susceptibles de faire l'objet d'un suivi dans l'enquête PISA et d'être inclus dans la série d'indicateurs de l'OCDE sur l'éducation (voir la section « La base de connaissances générales : les recherches sur l'efficacité de l'éducation »).

Créer une base de données durable au service des recherches pertinentes pour l'action publique

Les données PISA n'intéressent pas que les professionnels de l'éducation, les responsables politiques et les chercheurs en pédagogie. Elles sont de plus en plus utilisées par les économistes et les chercheurs en sciences sociales qui étudient des thématiques plus vastes, par exemple l'impact du capital humain sur la croissance économique (Hanushek et Woessmann, 2009) ou les variables prédictives de la réussite de l'intégration de familles de migrants (Stanat et Christensen, 2006).

Élargir la portée de l'enquête PISA pour en faire une base de données au service de recherches pertinentes pour l'action publique passe par une opérationnalisation très sophistiquée de *constructs* généraux tels que le milieu socio-économique des élèves et leur ascendance allochtone. Cela implique également la collecte d'autres informations (sur les valeurs civiques ou la santé, par exemple). Lors de la conceptualisation du contenu des questionnaires PISA, il faut tenir compte du fait que la base de données générée par l'enquête PISA doit répondre à des besoins – tant en matière de recherche que d'élaboration des politiques – qui vont au-delà du système d'indicateurs d'usage jusqu'ici.

À long terme, l'un des avantages majeurs de la base de données PISA sera de fournir des informations tendanciennes. La production de l'éducation est difficile à comprendre à partir de données transversales, car il est pratiquement impossible d'inclure toutes les variables pertinentes – y compris les données sur les performances antérieures des élèves. Toutefois, l'évolution de la performance à l'échelle des pays (voir Gustafsson, 2007 ; Hanushek et Woessmann, 2010), voire des établissements d'enseignement (voir la section « Extension de l'échantillon pour étudier la variation au niveau Établissement ») et des élèves (voir la section « Extensions longitudinales de niveau Élève ») peut être interprétée et



expliquée de façon probante si l'évolution des intrants et des processus est connue. Pour y parvenir, l'enquête PISA doit définir une série de variables qui resteront constantes pendant plusieurs cycles à l'avenir (voir la section « Répartition des contenus entre les cycles PISA : une structure globale de variables générales, les variables spécifiques au domaine d'évaluation et les variables d'extensions thématiques »).

Priorité à l'équité

L'équité se définit comme la répartition égale des ressources pédagogiques, des possibilités d'apprentissage et des résultats de l'éducation entre tous les individus scolarisés, quel que soit leur profil (OCDE, 2005a, p. 14). Étudier la variation des résultats cognitifs et non cognitifs entre les établissements d'enseignement et entre les élèves au sein des établissements est une première étape sur la voie de l'analyse de l'équité. L'écart entre les sujets très performants et les sujets peu performants est également intéressant. Il est possible de comparer des groupes d'élèves en fonction de leur établissement d'enseignement ou de leurs caractéristiques familiales. Ces variables ont souvent trait à la situation géographique (les régions ou le milieu rural ou urbain), aux effectifs des établissements d'enseignement publics ou privés, au milieu socio-économique, au sexe, à l'ascendance allochtone et à la langue maternelle. Des écarts de performance tenus entre des groupes d'élèves ainsi constitués peuvent dénoter une plus grande équité. L'équité peut aussi s'évaluer en fonction de l'accès de ces groupes à l'éducation, aux ressources pédagogiques et aux possibilités d'apprentissage. Les politiques en faveur de l'amélioration du niveau d'équité peuvent consister soit à proposer une répartition égale des ressources, soit à fournir un soutien supplémentaire aux groupes défavorisés. Enfin, une notion « méritocratique » de la justice admet que la diversité existe parmi les élèves, mais veut que des différences d'accès aux possibilités d'apprentissage ou dans la répartition des résultats de l'éducation ne soient pas dues à l'origine sociale (Cleary, 1968). L'enquête PISA fournit des données qui aident à évaluer les systèmes d'éducation en fonction de ces différents critères d'équité.

Les données du cycle PISA 2012 permettent non seulement d'évaluer le degré d'équité, mais également d'analyser des caractéristiques des systèmes d'éducation et des établissements d'enseignement en rapport avec l'équité, par exemple les politiques d'éducation. À l'échelle des systèmes, cela peut consister à comparer les systèmes en fonction de pratiques en matière d'orientation des élèves et du financement et de la gestion des établissements d'enseignement. Les politiques qui promeuvent l'équité peuvent consister à répartir les ressources de manière uniforme, à fournir un soutien ciblé aux établissements d'enseignement dont les effectifs sont défavorisés ou encore à réglementer les cours de soutien, à encadrer l'apprentissage de la langue d'enseignement ou à proposer d'autres activités extrascolaires. À l'échelle des établissements d'enseignement, les pratiques de regroupement par aptitudes et celles qui visent à aider les élèves ayant des besoins spécifiques d'éducation semblent pertinentes. La pédagogie différenciée et un environnement d'apprentissage structuré et propice peuvent avoir un impact sur l'équité à l'échelle des classes. De plus, la prévalence de ces caractéristiques peut être comparée entre les systèmes plus ou moins équitables.

L'enquête PISA inclut des variables d'équité jugées essentielles depuis son premier cycle, en 2000. Il est donc possible de suivre l'évolution du degré d'équité et d'identifier les changements qui y sont associés dans les politiques d'éducation et les processus agrégés. L'évolution de l'équité peut donc être comparée entre les pays et entre des groupes au sein même de ceux-ci sur une période de plus de 10 ans.

LA BASE DE CONNAISSANCES GÉNÉRALES : LA RECHERCHE SUR L'EFFICIENCE DE L'ÉDUCATION

Un certain nombre de résultats de l'éducation et de variables prédictives

L'école intègre les jeunes issus de milieux différents dans la société en leur permettant de participer à des activités cognitives et culturelles, de se livrer à des interactions sociales et de prendre part à la vie civique en général. Les établissements d'enseignement remplissent de nombreuses fonctions dans les sociétés modernes, notamment :

- instruire les jeunes et les préparer à réussir dans la vie, y compris dans leurs études et leur carrière professionnelle ;
- délivrer des diplômes et autres certificats pour guider les élèves dans la voie qui leur convient, que ce soit en termes d'études ou de carrière professionnelle, compte tenu de leur profil de compétence et de leurs objectifs personnels ; et
- cultiver le cœur et l'esprit des élèves, car l'école est un lieu de règles et de normes – dont certaines sont explicites, et d'autres informelles.

Dans l'enquête PISA, la littératie se définit comme la capacité des élèves à utiliser différents types de textes et d'outils mathématiques, et à se livrer à un raisonnement scientifique lorsqu'ils abordent, interprètent et résolvent des problèmes, et qu'ils prennent des décisions dans des situations de la vie réelle. Dans les sociétés modernes de la connaissance, la littératie est l'un des objectifs et des résultats les plus importants de l'éducation. L'enquête PISA cherche aussi à



évaluer des compétences cognitives plus transversales, en rapport par exemple avec les stratégies d'apprentissage, le raisonnement et la résolution de problèmes. La réussite à l'école – et dans la vie – dépend également de l'engagement des élèves, de leurs valeurs et de leurs convictions, et de la mesure dans laquelle ils respectent et comprennent l'autre, ils sont motivés à l'idée d'apprendre et de collaborer, et ils sont capables de gérer leur comportement d'apprentissage. Ces *constructs* peuvent être considérés comme des conditions essentielles à l'apprentissage cognitif, mais aussi, en soi, comme des objectifs de l'éducation, comme l'explique la publication de l'OCDE *Defining and Selecting Key Competencies* (DeSeCo) (Rychen et Salganik, 2003). C'est la raison pour laquelle l'enquête PISA s'intéresse à des résultats non cognitifs tels que les attitudes, les convictions, la motivation, les attentes et le comportement en matière d'apprentissage (autorégulation, stratégies et temps investi). L'absentéisme est une autre variable non cognitive qui suscite beaucoup d'intérêt, car elle constitue un indicateur (négatif) important de la mesure dans laquelle les élèves utilisent les possibilités d'apprentissage qui leur sont offertes et est un facteur prédictif probant de l'abandon scolaire et des comportements déviants (Kearney, 2008 ; Lee et Burkam, 2003). Ces résultats non cognitifs s'évaluent essentiellement sur la base du questionnaire « Élève », mais également sur la base du questionnaire « Établissement ». Des questionnaires proposés à titre d'option internationale, par exemple ceux sur le parcours scolaire et sur la maîtrise des technologies de l'information et de la communication, étendent ces indicateurs à des domaines qui revêtent un intérêt particulier pour un certain nombre de pays.

Une longue section des questionnaires « Élève » et « Établissement », ainsi que des questionnaires proposés à titre d'option internationale, concerne des facteurs contextuels en rapport avec des résultats cognitifs et non cognitifs. Ces éléments sont utilisés pour évaluer l'efficacité et pour définir des indicateurs. Ces facteurs font référence dans l'ensemble à des intrants ou à des processus. Les facteurs d'intrant concernent essentiellement le profil social et personnel des individus. Des caractéristiques structurelles telles que la taille des établissements d'enseignement et leur financement sont également considérées comme des intrants. Côté processus, l'apprentissage et l'enseignement sont considérés comme des processus majeurs, dont les aspects qualitatifs et quantitatifs sont évalués *via* plusieurs variables. De plus, les activités professionnelles des enseignants et des chefs d'établissement, et les politiques et pratiques scolaires, sont classées parmi les facteurs de processus.

Un problème particulier se pose dans les enquêtes internationales : les facteurs contextuels influant sur l'apprentissage des élèves sont à l'œuvre à quatre niveaux différents, soit les élèves et leur famille, leur classe, leur établissement d'enseignement et leur pays. Le processus de production se déroule comme suit :

- Des facteurs contextuels et structurels (intrants) *via* des interactions, des activités et des politiques (processus) jusqu'aux résultats cognitifs et non cognitifs : la performance des élèves en mathématiques, par exemple, dépend selon toute vraisemblance des activités au cours de mathématiques, qui dépendent à leur tour de la compétence de l'enseignant, de la taille de la classe, et des applications technologiques et autres ressources disponibles.
- Des facteurs d'ordre supérieur (système, établissement d'enseignement) à des facteurs propres à la classe et à des processus et des résultats propres aux élèves : la qualité de la direction pédagogique et de la collaboration entre enseignants dépend en partie du pouvoir de décision confié à l'établissement d'enseignement ou à un niveau administratif supérieur, ce qui façonne les stratégies pédagogiques et, donc, les activités d'apprentissage des élèves.

La structure fondamentale du modèle Intrant-Processus-Extrant a été élaborée dans les années 60 pour le compte de l'International Association for the Evaluation of Educational Achievement (IEA) (Purves, 1987). Une version plus récente de ce modèle, présentée dans le tableau 6.1, reprend ces dimensions. La première colonne affiche les quatre niveaux : Élève, Classe, Établissement d'enseignement et Pays. Les trois phases de production sont indiquées dans les autres colonnes, « Intrants », « Processus » et « Extrants ». Chaque cellule propose plusieurs *constructs* à titre d'exemple. Le modèle est relativement complet : il inclut la plupart des facteurs analysés dans l'enquête PISA et la quasi-totalité des *constructs* évoqués dans le présent cadre conceptuel. Il s'agit d'une sélection parmi les nombreux *constructs* abordés dans la littérature sur l'efficacité de l'éducation, mais le tableau reste trop fourni sous l'angle de la recherche. Le nombre de *constructs* qui pourront être pris en considération dépend du temps disponible pour le questionnaire et du système retenu en matière de rotation des items.

Côté intrants, certains facteurs sont relativement stables et difficiles à modifier, alors que d'autres peuvent être façonnés par des activités de développement scolaire ou des réorientations de l'action publique. Les processus sont généralement plus malléables, du moins indirectement (au travers de la formation initiale des enseignants et de leur développement professionnel). Les extrants sont le fruit des intrants et des processus. Toutefois, les limites entre les trois types de variables sont assez floues. Les extrants d'un cadre d'apprentissage deviennent les intrants d'un autre cadre, tandis que certains processus (les stratégies d'apprentissage) peuvent tout à fait être traités comme des intrants ou des extrants, selon des considérations théoriques, la nature des recherches ou des raisons pratiques.



Une fois inclus dans un modèle dynamique sur l'efficacité de l'éducation (voir Creemers et Kyriakides, 2008), les extrants deviennent des intrants au service du développement ultérieur. L'anxiété en mathématiques peut, par exemple, être un extrant et un intrant de l'éducation, qui a un impact sur les devoirs à domicile. De plus, les intrants peuvent s'influencer les uns les autres. Ainsi, dans de nombreux systèmes d'éducation, le niveau socio-économique des effectifs des établissements d'enseignement est en corrélation avec le financement des établissements, avec l'implication des parents dans la vie des établissements et même avec la qualité des enseignants. Tous ces facteurs se conjuguent pour offrir d'autres cadres d'enseignement et d'apprentissage (de meilleure qualité), et attirer des élèves issus de milieux favorisés (ou, plutôt, dont les parents sont plus favorisés), de sorte que la stratification sociale, les ressources et la qualité des processus se mêlent et sont difficiles à isoler.

Pour comprendre ces relations et effets correctement, et élaborer des modèles analytiques sophistiqués et adéquats, nous devons nous référer à la littérature sur l'éducation. Deux domaines de recherche sont particulièrement importants : les études sur l'efficacité de l'éducation en général (voir ci-dessous) et sur l'enseignement et l'apprentissage en particulier (voir la section « Les conditions d'apprentissage au service de la culture mathématique »).

Le tableau 6.1 décrit une grande série de facteurs aux différents niveaux qui constituent le système d'éducation. Leurs interactions sont complexes et difficiles à comprendre, et chaque niveau contribue dans une grande mesure à façonner la politique d'éducation et ses résultats probables.

Tableau 6.1

Taxinomie bidimensionnelle des résultats de l'éducation et de leurs variables prédictives

	Intrants	Processus	Extrants
Élèves	Sexe, année d'études, statut socio-économique. Parcours scolaire, notes. Ascendance allochtone, environnement et soutien familiaux. Expérience, attitudes et compétences en informatique. Ouverture, styles de résolution de problèmes.	Assiduité/absentéisme. Activités extrascolaires (participation à des activités extrascolaires, par exemple). Motivation, engagement. Stratégies d'apprentissage et de raisonnement, stratégies lors de tests. Temps d'apprentissage (y compris temps consacré aux devoirs et aux cours particuliers).	Performance en mathématiques. Attitudes, motivations et convictions à l'égard des mathématiques. Attitudes et comportements à l'égard de l'école en général (engagement, absentéisme, etc.). Motivation à l'idée d'apprendre, attentes scolaires.
Classes	Taille des classes, milieu socio-économique et composition ethnique des effectifs. Formation initiale et continue, et expertise des enseignants.	Qualité de l'instruction : structure, soutien, défis. Possibilité d'apprendre : programme de cours, tâches à accomplir, activités en rapport avec les mathématiques. Temps d'enseignement, pratiques de regroupement, évaluation et feed-back.	Variables agrégées au niveau Élève.
Établissements d'enseignement	Milieu socio-économique et composition ethnique des effectifs. Milieu socio-économique de la communauté. Financement public ou privé des établissements. Taille des établissements. Implications des parents.	Importance de la performance, normes partagées, leadership, moral des enseignants, coopération entre les enseignants, développement professionnel. Politiques d'admission et de recrutement, orientation des élèves, offre de cours, programme de cours, évaluation. Relations entre élèves et enseignants, environnement favorable.	Variables agrégées au niveau Élève. Passage de classe, redoublement et taux d'obtention d'un diplôme. Assiduité.
Pays (systèmes)	Richesse, (in-)égalité sociale. Politiques en matière de diversité.	Financement des établissements d'enseignement, politiques de différenciation et d'orientation des élèves, politiques de développement professionnel des enseignants, soutien aux élèves ayant des besoins spécifiques d'éducation ou appartenant à un groupe linguistique minoritaire, politiques de recrutement et de certification. Politiques de responsabilisation et d'évaluation, instances décisionnelles.	Variables agrégées au niveau Élève. Taux moyen d'obtention d'un diplôme.



La recherche sur l'efficacité de l'éducation vise essentiellement à identifier « les facteurs en rapport avec l'enseignement, le programme de cours et le cadre d'apprentissage, tels que la classe, l'établissement d'enseignement et les niveaux supérieurs, qui peuvent expliquer, directement ou indirectement, les écarts de performance entre les élèves, compte tenu de leurs caractéristiques, par exemple leurs aptitudes, leur milieu socio-économique et leur performance antérieure » (Creemers et Kyriakides, 2008, p. 12). Toutefois, les pratiques ne présentent pas la même efficacité pour tous les élèves d'un établissement d'enseignement, ni pour tous les systèmes d'éducation, contextes locaux ou établissements d'enseignement. De plus, différentes conclusions peuvent être tirées en fonction du type de résultat ciblé (Kyriakides et Tsangaridou, 2004). Dans ce contexte, les études récentes sur l'efficacité de l'éducation tiennent également compte des interactions entre les intrants et analysent les différentiels d'efficacité et les pratiques différenciées. De nombreuses études ont été menées dans ce domaine (voir, par exemple, Creemers et Kyriakides, 2008 ; Scheerens, 2000 ; Teddlie et Reynolds, 2000). Les sections suivantes résument les résultats obtenus pour chaque cellule de la taxinomie présentée dans le tableau 6.1. Par ailleurs, des façons d'opérationnaliser les grands constructs et les considérations à retenir en matière de conception seront présentées à la fin de chaque section.

Intrants et processus à l'échelle des systèmes

Dans la plupart des pays, les décisions en rapport avec la politique d'éducation et l'affectation des ressources sont du ressort d'une entité fédérale ou nationale. Dans les systèmes plus centralisés, l'instance suprême de décision est souvent le ministère de l'Éducation, alors que dans les systèmes moins centralisés, le pouvoir de décision peut être partagé avec les autorités des entités fédérées ou avec des agences régionales, au niveau des districts. Les facteurs systémiques sont classés parmi les intrants, les processus et les extrants dans cette taxinomie, mais leur rôle dans les modèles analytiques sera très vraisemblablement celui de variables modératrices ; en d'autres termes, les facteurs systémiques ont un impact sur les relations entre les intrants, les processus et les extrants au niveau inférieur.

Au niveau du système, les intrants de base concernent les ressources matérielles et leur répartition au sein des pays. Selon certaines études, le revenu par habitant est une variable prédictive probante de la performance des élèves tant entre les pays qu'au sein même de ceux-ci (Baker *et al.*, 2002). La richesse permet de mobiliser des ressources au service de la performance des élèves (Baumert *et al.*, 2005). Parallèlement, les résultats scolaires peuvent aussi influencer sur la prospérité économique des pays. Les inégalités sociales sont en corrélation avec des inégalités dans l'éducation, à cause de la variation des ressources familiales et des conditions d'apprentissage. De plus, certains pays prévoient un financement similaire par élève, alors que d'autres accordent un soutien spécifique aux élèves qui vivent en zone rurale, qui sont défavorisés, qui sont issus de l'immigration ou qui souffrent de difficultés d'apprentissage ou de troubles physiques. Les orientations d'un pays concernant l'investissement dans l'éducation peuvent avoir des implications à la fois sur le niveau de performance ainsi que sur l'équité des résultats et de l'accès aux carrières. Il est un autre aspect important en rapport étroit avec les inégalités sociales : les politiques en matière d'immigration et de diversité (Stanat et Christensen, 2006). Les élèves issus de l'immigration accusent des résultats inférieurs dans plusieurs des matières principales dans de nombreux pays, alors qu'ils font jeu égal avec les autres élèves, voire les devancent dans d'autres pays (OCDE, 2012b). Les écarts de performance sont plus réduits dans les pays qui recrutent activement des immigrants très instruits. En revanche, des pourcentages élevés d'immigrants défavorisés peuvent poser des problèmes particuliers aux systèmes d'éducation. Des recherches montrent aussi que des politiques de soutien linguistique permettent de réduire les désavantages des élèves dont la langue maternelle n'est pas la langue d'enseignement.

Pour décrire les systèmes scolaires, il est important aussi d'étudier les politiques concernant le regroupement par aptitudes et la responsabilité de la gestion des établissements d'enseignement. Dans certains pays, les élèves sont répartis entre les établissements en fonction de filières d'enseignement (générale, professionnelle ou technique) selon leurs résultats scolaires et/ou leurs préférences. Dans ces systèmes différenciés, les élèves n'ont pas tous accès aux mêmes possibilités d'apprentissage. Ces pratiques reviennent régulièrement dans les débats sur la question de savoir si la différenciation prive de possibilités d'apprentissage des élèves dont le développement intellectuel est plus tardif ou qui sont issus de milieux socio-économiques moins favorisés (Levin, 1978). Les systèmes inclusifs, où tous les élèves suivent au moins un tronc commun minimal, peuvent être plus efficaces en termes d'équité (OCDE, 2007b). Toutefois, même au sein des établissements d'enseignement, des classes peuvent être constituées d'effectifs hétérogènes ou regroupés par aptitudes. Les pays se distinguent non seulement en fonction de leur degré de différenciation, mais également en fonction de l'âge auquel les élèves sont répartis entre les établissements d'enseignement ou entre les filières, et de la possibilité de changer de filière. Enfin, point important s'il en est, les politiques d'admission et de passage de classe ont un impact sur les résultats aux épreuves PISA ; en effet, la performance augmente si les élèves progressent plus vite entre les années d'études – toutes choses étant égales par ailleurs (Gustafsson, 2007).



Il ressort des cycles antérieurs de l'enquête PISA que le financement et la gestion des établissements d'enseignement sont le plus souvent publics, mais que dans une majorité des pays participants, il existe aussi un certain nombre d'établissements d'enseignement dont le financement et la gestion sont essentiellement du ressort d'organisations non gouvernementales, par exemple des associations confessionnelles ou des entreprises privées. Il existe aussi des établissements d'enseignement gérés par des organisations privées, mais essentiellement financés par les pouvoirs publics, ce que l'on appelle « les établissements d'enseignement privés subventionnés par l'État ». Les deux types d'établissements d'enseignement s'inscrivent dans le même cadre réglementaire, mais les établissements privés jouissent généralement d'une plus grande latitude pour prendre des décisions en matière d'éducation. Ils représentent donc une forme de décentralisation de la prise de décision. Le choix entre les établissements publics et privés est considéré comme un mécanisme qui encourage les établissements à se concurrencer pour attirer les élèves, ce qui aurait pour effet d'améliorer le niveau de performance global du système d'éducation (Belfield et Levin, 2002). Toutefois, des résultats antérieurs de l'enquête PISA suggèrent que la performance moyenne plus élevée des établissements d'enseignement privés indépendants et subventionnés par l'État s'explique essentiellement par le fait que leurs élèves sont issus de milieux socio-économiques favorisés, ce qui a des conséquences en termes de composition des effectifs (OCDE, 2007b). Le pourcentage d'établissements d'enseignement privés et leurs caractéristiques peuvent donc influencer sur l'équité, comme le degré de différenciation des élèves (OCDE, 2012c).

Depuis quelques années, de nombreux pays se demandent comment accroître la réactivité de leurs établissements d'enseignement et leur faculté d'adaptation aux besoins d'élèves, de groupes d'élèves et de communautés spécifiques. Beaucoup prétendent que lorsque la prise de décision est du ressort d'une instance centrale, le gouvernement national par exemple, le recrutement et la formation des enseignants et les programmes de cours sont trop rigides pour répondre aux besoins d'établissements d'enseignement dont les effectifs d'élèves sont différents. C'est la raison pour laquelle de nombreux pays ont entrepris de décentraliser leur système d'éducation et de déléguer le pouvoir de décision à des niveaux plus proches des élèves et des communautés, dans l'espoir d'améliorer la performance de leurs établissements d'enseignement (Hannaway et Carnoy, 1993). Toutefois, les études sur la relation entre cet aspect et les résultats scolaires n'ont pas permis de recueillir des éléments probants à l'appui d'une forte corrélation (voir, par exemple, Schmidt et McKnight, 1998). Certaines analyses secondaires des résultats des élèves à des épreuves internationales suggèrent que l'autonomie scolaire est bénéfique si elle va de pair avec une forme d'évaluation descendante ou l'administration d'épreuves centralisées (Woessmann, 2006).

Il peut être important de soumettre les établissements d'enseignement et les élèves à des évaluations pour garantir que les élèves atteignent tous des normes supérieures et sont tous préparés à poursuivre des études, en particulier dans les systèmes décentralisés (Carnoy *et al.*, 2003). Plusieurs pays ont choisi cette voie, mais leurs approches en matière d'évaluation et de responsabilisation sont sensiblement différentes. Certains pays ont instauré un système périodique d'évaluation pour recueillir des données normalisées sur la performance des élèves et des établissements d'enseignement à l'échelle fédérale, nationale ou régionale. D'autres pays imposent ce type d'évaluations, mais laissent aux autorités régionales et locales le soin de concevoir et d'administrer les épreuves. D'autres encore ne procèdent pas à des évaluations systématiques. Par ailleurs, la responsabilisation s'étend aussi aux conséquences des résultats d'évaluation (Carnoy *et al.*, 2003 ; Koretz, 2008). Dans certains pays, les tests et les enquêtes servent uniquement à informer l'opinion et les établissements d'enseignement. Les résultats aux tests peuvent pourtant être utilisés à d'autres fins, par exemple encourager l'augmentation de la performance ou identifier les points faibles des établissements et les aider à y remédier. Des systèmes différents peuvent avoir des conséquences différentes, que ce soit en termes d'incitations à apprendre, d'uniformité de l'instruction ou de résultats scolaires.

Outre ces variables systémiques plutôt descriptives, un certain nombre de politiques et pratiques plus spécifiques ont été proposées pour être incluses dans le modèle dynamique d'efficacité de l'éducation (Creemers et Kyriakides, 2008). Ces politiques et pratiques visent à aider les établissements d'enseignement à créer un environnement d'apprentissage plus positif, à dispenser un enseignement aux normes qualitatives et quantitatives élevées, et à offrir aux élèves un large éventail de possibilités d'apprentissage. Peu d'études ont intégré ces processus systémiques. Le cycle PISA 2012 permet, dans une certaine mesure, d'analyser leur prévalence et leur corrélation avec la performance dans un échantillon important de pays. De plus, l'existence de ces politiques et la nature de leur mise en œuvre pourront être étudiées.

Les possibilités d'apprentissage dépendent des programmes de cours nationaux. Le programme de l'IEA *Trends in Mathematics and Science Study* (TIMSS) a identifié une variation sensible entre les programmes de cours tels que prévus et ceux appliqués dans les pays, en particulier en fonction du temps investi, de la cohérence et des priorités, ce qui explique en partie les écarts de performance observés entre les élèves, du moins dans des évaluations basées sur les programmes de cours telles que l'étude TIMSS (Schmidt *et al.*, 2001 ; Schmidt et Houang, 2007 ; voir également la



section « Possibilités d'apprentissage et qualité de l'instruction : évaluation de l'environnement d'apprentissage ». Les responsables politiques peuvent aussi soutenir et encourager la création de possibilités extrascolaires d'apprentissage, en plus de l'apprentissage formel. Ils peuvent, par exemple, lancer des concours de mathématiques ou fournir aux établissements d'enseignement les moyens et le soutien requis pour organiser des programmes extrascolaires. Par ailleurs, les pays peuvent choisir des orientations politiques différentes au sujet des cours payants et d'autres formes d'« éducation parallèle » (Baker et LeTendre, 2005).

La qualité de l'instruction peut être contrôlée à l'échelle du système au travers de la définition de normes d'enseignement, dont le respect est vérifié. Cela implique la formulation d'attentes explicites sur la base des recherches les plus récentes et présuppose de surcroît que les évaluations ne portent pas uniquement sur les extraits, mais également sur l'enseignement et l'instruction. Les responsables politiques peuvent concevoir des politiques et mobiliser des ressources au service de la formation continue pour contribuer à la professionnalisation du personnel des établissements d'enseignement. Par activités de développement professionnel, on entend les activités visant à développer les connaissances, les compétences, l'expertise et d'autres caractéristiques liées au corps enseignant (OCDE, 2009b, p. 49). Dans de nombreux pays, les enseignants sont dans l'obligation de consacrer un certain nombre de jours à des activités de développement professionnel. Toutefois, le niveau et les tendances de participation aux différents types d'activités de développement professionnel varient sensiblement entre les systèmes d'éducation. La participation peut être encouragée via l'octroi de primes ou de compléments de salaire et l'adaptation des emplois du temps. Toutefois, les résultats de l'Enquête internationale de l'OCDE sur les enseignants, l'enseignement et l'apprentissage (*Teaching and Learning International Survey*, TALIS) donnent à penser que l'effet de ces mesures est limité (OCDE, 2009b).

Enfin, des valeurs et des normes culturelles influent également sur le comportement des différentes parties prenantes. L'engagement des élèves et de leurs parents à l'égard de l'établissement d'enseignement et de l'apprentissage dépend de la mesure dans laquelle ils apprécient les valeurs en rapport avec l'éducation. Le statut global des enseignants est important, car il a un impact sur l'attractivité de la profession aux yeux des personnes qui l'envisagent. L'implication de la communauté dans la scolarisation peut améliorer la sensibilité via l'adaptation aux besoins des populations locales.

Concernant l'élaboration des variables systémiques, certaines, telles que la richesse, l'investissement dans l'éducation et l'équité de sa répartition, la composition des populations de migrants et les règles en matière de différenciation et d'admission, peuvent être dérivées de bases de données existantes, par exemple du système d'indicateurs de l'OCDE sur l'économie, le bien-être et l'éducation⁵. D'autres pourraient être dérivées d'un questionnaire sur le système à soumettre à la direction de l'administration qui traiterait, par exemple, des politiques en rapport avec le recrutement et le développement professionnel des enseignants, l'évaluation et la responsabilisation, le financement des établissements, le degré de centralisation ou de décentralisation, les normes et les activités extrascolaires, et le soutien aux élèves ayant des besoins spécifiques d'éducation ou appartenant à des minorités. Plusieurs indicateurs peuvent être dérivés des données PISA agrégées, par exemple l'indicateur relatif à la participation des établissements d'enseignement au processus de prise de décision calculé à partir du questionnaire « Établissement ». Des données sur les élèves, les enseignants et les parents peuvent être agrégées à l'échelle nationale pour fournir des indicateurs culturels, par exemple la valeur accordée à l'éducation par rapport à d'autres aspects de la vie personnelle et sociale.

Intrants et processus à l'échelle des établissements d'enseignement

L'organisation des établissements d'enseignement est complexe et varie considérablement, non seulement entre les pays, mais également au sein même des pays, c'est-à-dire entre les entités fédérées, les provinces, les régions, les districts scolaires et les établissements d'enseignement. Parmi les intrants des établissements d'enseignement, ceux en rapport avec leur effectif d'élèves sont susceptibles d'influer sur les résultats scolaires. Les établissements d'enseignement qui accueillent des élèves issus de milieux socio-économiques plus favorisés tendent à créer des environnements propices à l'apprentissage académique où les attentes sont élevées. De plus, ces établissements peuvent attirer de meilleurs enseignants. Cela avantage même les élèves issus de milieux socio-économiques moins favorisés qui les fréquentent. Les intrants relatifs aux élèves dépendent vraisemblablement de la richesse de la communauté environnante, ce qui a des effets sur l'apprentissage des élèves. Les quartiers plus riches tendent à craindre moins que les autres les effets possibles de la criminalité et de la violence, et à disposer de davantage de ressources pour accompagner les établissements d'enseignement de façon informelle. Par ailleurs, les bons établissements d'enseignement peuvent attirer des parents engagés, ce qui a également un impact sur la communauté. Le pourcentage d'élèves dont la langue maternelle n'est pas la langue d'enseignement est une dimension quelque peu différente. La diversité linguistique est une richesse pour les établissements d'enseignement, mais elle a aussi un coût, dans la mesure où elle implique la mise en place de systèmes spécifiques de soutien et peut nécessiter l'intervention d'enseignants spécialisés pour remédier aux difficultés



de compréhension et tenir compte de différences culturelles potentielles. Dans de nombreux pays, le milieu socio-économique varie fortement entre les communautés urbaines, suburbaines et rurales.

Lorsque les établissements d'enseignement encouragent et organisent l'implication des parents, ceux-ci peuvent être plus efficaces pour aider leurs enfants à progresser et pour apporter leur soutien à l'établissement. De plus, les parents qui participent à des activités scolaires sont plus susceptibles de se porter volontaires pour aider les établissements, ce qui accroît les ressources disponibles. Les parents qui connaissent les attentes des établissements d'enseignement parce qu'ils sont familiarisés avec les programmes et les exigences des enseignants sont plus à même d'aider leurs enfants. Enfin, amener les parents à participer activement aux initiatives éducatives et scolaires permet de créer des réseaux sociaux qui leur offrent la possibilité de faire connaissance et de s'entraider. Ces réseaux sociaux entraînent vraisemblablement une augmentation de la performance globale via la constitution d'un « capital social » (Coleman, 1988).

On estime également que la taille des établissements d'enseignement est en corrélation avec la performance. Les établissements qui accueillent beaucoup d'élèves peuvent leur offrir un éventail plus diversifié d'enseignants, de cours, de matières à option et d'activités extrascolaires, ce qui permet aux élèves de choisir les cours qui leur conviennent le mieux sur le plan éducatif et qui les motivent le plus. Les établissements plus grands tendent toutefois à être plus impersonnels, et le soutien aux élèves peut y être moins individualisé. Selon certaines études, la performance et l'engagement des élèves âgés de 15 ans sont plus élevés et moins variables dans les établissements d'enseignement secondaire de plus petite taille (Coleman, 1988). La question de la taille optimale des établissements d'enseignement reste sans réponse. En fait, certaines études suggèrent que la taille des établissements a des effets qui varient selon les groupes d'élèves, par exemple ceux constitués en fonction du milieu socio-économique (Lee et Smith, 1997).

En plus de ces intrants, un certain nombre de processus sont des variables prédictives directes ou indirectes des résultats à l'échelle des établissements d'enseignement. La caractéristique malléable la plus importante des établissements d'enseignement est vraisemblablement le climat (ou la qualité de l'environnement) scolaire. Le climat scolaire dépend non seulement des normes et des valeurs, mais également de la qualité des relations et de l'ambiance générale. Les priorités académiques – un consensus général au sujet de la mission de l'établissement d'enseignement et de la valeur accordée à l'éducation – favorisent l'apprentissage. De plus, un environnement d'apprentissage discipliné optimise l'exploitation du temps d'apprentissage. En revanche, les environnements qui se caractérisent par un manque de discipline et de respect sont contre-productifs, tant pour les enseignants que pour les élèves, et éloignent les établissements d'enseignement de leur mission. Les environnements disciplinés peuvent être façonnés par l'adoption de règles cohérentes (pour remédier à des comportements inappropriés des élèves, l'absentéisme par exemple). Les établissements d'enseignement efficaces sont capables de s'adapter aux besoins spécifiques des élèves et des enseignants. Ils organisent des cours de soutien ou de perfectionnement pour les élèves ayant des besoins spécifiques d'éducation, les élèves très doués et les élèves dont la langue maternelle n'est pas la langue d'enseignement. De plus, ils proposent des services d'orientation et d'aide aux devoirs. Plus généralement, la variété et la qualité des programmes extrascolaires peuvent aussi favoriser l'apprentissage des élèves et sont particulièrement pertinentes pour les résultats non cognitifs et les perspectives professionnelles des élèves.

Trois aspects de l'organisation des établissements d'enseignement sont pertinents pour améliorer l'environnement d'enseignement et d'apprentissage : la collaboration entre enseignants, la professionnalisation et la direction. La collaboration entre enseignants fournit à ceux-ci un soutien pratique et affectif, et contribue à leur professionnalisation. Elle se présente sous diverses formes : coordonner les pratiques, rendre les programmes de cours plus cohérents, et fournir un soutien et un feedback mutuels. Dans de nombreux pays, il est relativement courant que les enseignants s'échangent du matériel pédagogique ou discutent des difficultés d'apprentissage de certains élèves. Il existe des formes plus sophistiquées de coopération, par exemple les activités collectives d'apprentissage, telles que les visites d'études, la transmission de feedback, la participation à des activités de développement professionnel, l'organisation d'activités communes à plusieurs classes ou groupes d'âge (OCDE, 2009b). Ces pratiques sont particulièrement utiles pour transformer les établissements en structures d'apprentissage ainsi que pour fournir un feedback constructif aux enseignants et un soutien à la participation à des activités de développement professionnel qui répondent à leurs besoins spécifiques. Les chefs d'établissement doivent s'acquitter de tâches administratives en rapport avec la réglementation et les budgets, gérer les infrastructures et s'occuper des relations publiques. Toutefois, leur mission consiste avant tout à s'assurer que l'instruction et l'apprentissage sont de grande qualité dans leur établissement. Depuis quelques décennies, le concept de direction pédagogique, qui met l'accent sur cet aspect, est abondamment traité dans la littérature (Blase et Blase, 1998 ; Hallinger et Heck, 1998 ; Heck et Hallinger, 2005 ; Krüger *et al.*, 2007 ; Leithwood et Riehl, 2005 ; Witziers *et al.*, 2003). Toutefois, les résultats empiriques à propos des relations entre la direction pédagogique et la performance des élèves ne sont pas cohérents, vraisemblablement parce que la direction pédagogique a un effet indirect qui subit l'influence d'autres processus à l'échelle de l'établissement et de la classe.



Le volume d'instruction peut être déterminé par les mesures politiques et la réglementation sur les cours, les horaires, la gestion du temps d'enseignement, l'assiduité des élèves et l'implication des parents. Des directives sur l'enseignement, les procédures de regroupement et le comportement des enseignants peuvent contribuer à améliorer la qualité de l'enseignement. Les possibilités d'apprentissage dépendent fortement des programmes de cours et du choix du matériel pédagogique. Toutefois, la description formelle des opérations à mener au sein des établissements d'enseignement est souvent trompeuse, car la mise en œuvre varie énormément (Fullan, 1992 ; Fullan et Stiegelbauer, 1991). C'est pourquoi les questionnaires PISA tentent d'obtenir des informations sur les opérations directement auprès des participants, en particulier des chefs d'établissement.

Il ne suffit pas de présenter ces politiques et pratiques, il faut aussi en évaluer les effets. L'évaluation des établissements d'enseignement peut être utilisée pour identifier les forces et les faiblesses des opérations, étudier les pratiques et analyser leur impact sur les résultats. L'enquête TALIS montre que les pratiques d'évaluation varient fortement entre les pays, non seulement quant à leur fréquence, mais également quant à leurs sources d'information, à leurs priorités et à leurs conséquences (OCDE, 2009b). Les effets des évaluations des établissements d'enseignement sur les résultats n'ont fait l'objet que de peu d'études à ce jour (Creemers et Kyriakides, 2008).

Le degré de centralisation (par opposition à l'autonomie) de la gouvernance est une caractéristique quelque peu différente qui a un impact sur la prise de décision. Dans certains pays, le recrutement des enseignants, l'élaboration des programmes de cours et l'évaluation des élèves s'effectuent au niveau central. Dans d'autres pays, les établissements d'enseignement doivent simplement se conformer aux orientations qui leur sont fournies au sujet des opérations scolaires de routine. À l'autre extrême du spectre, les établissements d'enseignement sont très autonomes : ils choisissent leurs enseignants, influent sur leur rémunération et sélectionnent la façon dont ils évaluent la performance de leurs élèves. Depuis quelques années, l'autonomie scolaire est souvent citée comme un moyen de réduire la rigidité bureaucratique ; elle pourrait avoir un impact sur les résultats de l'apprentissage (Bottani et Favre, 2001 ; Chubb et Moe, 1990). Toutefois, aucune conclusion définitive n'a encore été tirée, et la décentralisation fait toujours débat.

Concernant l'élaboration des variables scolaires, la plupart sont dérivées des réponses des chefs d'établissement à des questions sur les ressources de leur établissement (ou le manque de ressources), les programmes de cours (les horaires, les filières, les cours de soutien et de perfectionnement, les activités extrascolaires), le climat scolaire (les attentes, le moral des enseignants et des élèves, l'implication des parents et les problèmes de comportement) et les activités professionnelles (la collaboration entre enseignants, les normes, le leadership et les procédures d'évaluation). L'ajout de données dérivées d'un questionnaire soumis aux enseignants et agrégées à l'échelle des établissements d'enseignement permettrait de brosser un tableau plus vaste et probablement plus probant de la situation à ces égards. Cette approche est applicable dans les pays où le cycle PISA 2012 est associé au prochain cycle de l'enquête TALIS (OCDE, 2009b). En plus des informations dérivées du questionnaire « Établissement » sur ces sujets, des données de niveau Élève, dérivées notamment du questionnaire facultatif sur la maîtrise des TIC, et des données de niveau Parents peuvent être agrégées pour élaborer des indicateurs sur la composition des effectifs d'élèves, le capital social, les ressources des établissements d'enseignement et le climat scolaire (par exemple les TIC disponibles, les relations entre enseignants et élèves, l'attitude des parents à l'égard de l'établissement et leur implication dans la vie de l'établissement). Enfin, le questionnaire sur le parcours scolaire inclut notamment des questions sur les services d'orientation proposés par les établissements d'enseignement.

Le questionnaire « Établissement » administré lors du cycle PISA 2012 permet également de recueillir des données qui vont au-delà de la performance et de la conduite des élèves directement évalués. Les chefs d'établissement ont, par exemple, été invités à répondre à des questions sur l'absentéisme des élèves, ainsi que sur les taux de passage de classe et d'obtention d'un diplôme. Ces résultats comportementaux « objectifs » sont vraisemblablement plus comparables, du moins dans un même système d'éducation, que la plupart des indicateurs dérivés des déclarations ou des descriptions d'attitudes faites par les répondants. De plus, ils représentent des résultats de niveau Établissement, tandis que la performance théorique des élèves peut en grande partie être estimée sur la base de variables de niveau Enseignant et Classe.

Intrants et processus à l'échelle des classes

La plupart des expériences d'apprentissage que vivent les élèves se déroulent en classe. C'est en classe que les élèves sont exposés aux contenus, au matériel pédagogique, aux stratégies pédagogiques, à la composition des effectifs et au climat qui y règne. Comme des recherches antérieures l'ont montré, des variables proximales – dont des caractéristiques et pratiques propres à la classe – sont en corrélation plus forte avec la performance des élèves que des variables plus distales, telles que des facteurs à l'échelle des établissements d'enseignement et des systèmes (voir, par exemple, Wang *et al.*, 1993).



Les intrants les plus importants à l'échelle de la classe sont le contexte et l'enseignant (Wayne et Youngs, 2003). Par contexte, on entend les caractéristiques des élèves et leur regroupement, et la taille de classe. Les considérations à propos du milieu socio-économique et de la langue maternelle des élèves sont similaires à l'échelle des classes et des établissements. Dans de nombreux cas, la composition des effectifs varie non seulement entre les établissements d'enseignement, mais également au sein même de ceux-ci, en particulier lorsque les élèves sont répartis entre les classes en fonction de leur niveau de performance. Cette forme de regroupement a pour principe de permettre aux enseignants d'améliorer l'enseignement en l'adaptant aux aptitudes de leurs élèves, mais de nombreux éléments montrent qu'elle peut creuser les écarts de performance entre les groupes (Gamoran, 1992 ; Oakes, 2005). Les formes plus souples de regroupement par aptitudes, par exemple celles qui se limitent à certaines tâches ou à certaines périodes, ou le regroupement hétérogène, peuvent toutefois être utiles pour appliquer des pratiques pédagogiques différenciées et collaboratives (Slavin, 1990).

L'évaluation et la notation des élèves au jour le jour est un aspect spécifique des pratiques en classe, qui semble relativement important dans le contexte des études telles que l'enquête PISA. Les élèves sont-ils soumis régulièrement à des tests normalisés ? Reçoivent-ils un feedback des enseignants sous la forme de notes ou de commentaires écrits ou oraux sur leurs notes ? Les établissements d'enseignement, les enseignants, les élèves et/ou leurs parents reçoivent-ils des informations sur la performance des élèves par rapport à des normes générales ? Comment utilisent-ils ces informations ? Les élèves s'entraînent-ils à passer des tests ? Comment ces pratiques évoluent-elles dans le temps dans les différents pays ? Comme le montrent des études sur les évaluations formatives et sommatives (Brookhart, 2009) ainsi que sur les conditions et les effets de tests dont les enjeux sont très importants (Borko *et al.*, 2007 ; Koretz, 2008), les résultats des élèves à des épreuves normalisées telles que celles administrées dans l'enquête PISA peuvent varier selon les politiques et pratiques en matière d'évaluation.

On estime généralement que la taille des classes a un effet (négatif) sur la performance des élèves. Lorsque les effectifs sont moins nombreux par classe, les élèves ont davantage de possibilités de participer et il est plus facile pour les enseignants d'utiliser des pratiques pédagogiques ciblées sur des élèves ou des groupes d'élèves. De plus, les enseignants ont moins de copies à corriger, ce qui leur laisse plus de temps pour se concentrer sur le feedback et concevoir des activités et des évaluations plus riches. Des recherches internationales montrent pourtant que la réduction des effectifs doit être conséquente pour avoir quelque impact et qu'elle est moins pertinente dans l'enseignement secondaire que dans l'enseignement primaire (Gustafsson, 2007). Enfin, les effets de la taille des classes semblent spécifiques aux cultures. Les tailles de classes sont relativement grandes dans de nombreux pays asiatiques, où la performance moyenne des élèves est pourtant élevée.

Les enseignants se caractérisent par leur formation, leur expérience et leur expertise. Une formation académique poussée, une spécialisation dans la matière enseignée et une expérience professionnelle sont autant de qualifications souhaitables et d'indicateurs de la qualité des enseignants dans certains modèles. Toutefois, les éléments à l'appui de leur corrélation avec la performance des élèves ne sont pas cohérents (voir, par exemple, Hanushek et Rivkin, 2007 ; Libman, 2009 ; Mullis et Martin, 2007 ; Zuzovsky, 2009). Des effets plus importants sont établis avec l'expertise des enseignants (voir, par exemple, Baumert *et al.*, 2009 ; Hill, Rowan et Ball, 2005). Ces effets sont décrits comme issus de la combinaison des connaissances des enseignants en pédagogie générale et appliquée à leur matière, de leurs convictions, de leur motivation et de la capacité à s'autogérer (voir, par exemple, Bromme, 1997 ; Brunner *et al.*, 2006 ; Shulman, 1987). Outre les intrants décrits ci-dessus, un certain nombre de processus à l'œuvre à l'échelle de la classe améliorent, selon certaines études, l'efficacité de l'éducation. Il s'agit en particulier du climat et de la qualité de l'instruction en classe. Il ressort de certaines recherches que l'apprentissage des élèves est dans l'ensemble favorisé par un climat positif et respectueux, relativement peu perturbé et axé sur la performance des élèves (Creemers et Kyriakides, 2008 ; Harris et Chrispeels, 2006 ; Hopkins, 2005 ; Scheerens et Bosker, 1997).

Les aspects majeurs d'un climat positif en classe sont : des interactions propices entre enseignants et élèves, de bonnes relations entre enseignants et élèves, une priorité aux accomplissements et un environnement d'apprentissage qui se caractérise par des règles claires de discipline. La qualité de l'instruction est toutefois une dimension plus complexe. Des éléments suggèrent qu'il n'y a pas qu'une seule façon de bien enseigner. Des cours bien structurés et bien gérés au rythme étudié, un suivi rapproché, une présentation claire et un feedback encourageant et instructif (les aspects clés de l'« instruction directe ») sont en corrélation positive avec la performance des élèves. Toutes ces composantes contribuent à créer un environnement propice à l'apprentissage en classe et optimisent l'exploitation du temps d'apprentissage. Toutefois, la motivation et les résultats non cognitifs des élèves augmentent sous l'effet d'autres caractéristiques qualitatives de l'instruction, tels que le climat en classe et les relations entre enseignants et élèves qui favorisent l'autonomie des élèves, la compétence et les rapports sociaux. De plus, pour améliorer la compréhension



conceptuelle, l'instruction doit se baser sur des contenus exigeants (Brown, 1994). Par ailleurs, différents sous-groupes d'élèves peuvent bénéficier de pratiques différentes en matière d'instruction. Les enseignants doivent donc orchestrer les activités d'apprentissage pour répondre aux besoins spécifiques de leur classe. Klieme *et al.* (2009) ont condensé ces aspects dans un cadre de qualité de l'instruction à trois dimensions : *i*) une gestion claire et bien structurée de la classe ; *ii*) un climat propice et axé sur les élèves en classe ; et *iii*) une activation cognitive sur la base de contenus exigeants. Plusieurs études indépendantes sur les cours de mathématiques dans l'enseignement secondaire ont confirmé depuis lors cette structure triarchique de la qualité de la classe et corroborent dans une certaine mesure l'impact cognitif et motivationnel supposé (*TIMSS-Video* : Klieme *et al.*, 2001 ; *COACTIV* : Baumert *et al.*, 2009 ; *Pythagoras* : Lipowsky *et al.*, 2009). Klieme et Rakoczy (2003, voir également Kunter *et al.*, 2008) ont identifié des structures similaires dans des extensions nationales de l'enquête PISA. Toutefois, la qualité de l'instruction, en particulier l'activation cognitive, est dans une grande mesure propre aux matières. Des aspects spécifiques aux mathématiques sont décrits dans la section « Les conditions d'apprentissage au service de la culture mathématique ».

Concernant l'élaboration des variables, l'enquête PISA n'inclut pas de niveau Classe dans la mesure où les échantillons d'élèves âgés de 15 ans sont prélevés dans toutes les classes des établissements d'enseignement échantillonnés, si ce n'est dans les quelques pays qui ont étendu leur échantillon pour y inclure des composantes basées sur les années d'études. Cette conception ne va pas sans difficultés pour analyser des effets spécifiques aux classes et aux enseignants (Opdenakker et van Damme, 2000 ; Van Landeghem *et al.*, 2005), et étudier des modèles bien spécifiés d'efficacité de l'éducation. De plus, elle renforce l'orientation de l'enquête PISA sur le rendement. L'enquête PISA peut toutefois rendre compte des intrants et des processus à l'échelle de la classe de façon descriptive, compte tenu de variables dont d'autres sources établissent le lien avec la performance des élèves et moyennant leur évaluation *via* les questionnaires « Élève », « Établissement » et, le cas échéant, « Enseignant ». Les chefs d'établissement ont, par exemple, répondu à des questions au sujet de caractéristiques contextuelles des enseignants et de leurs orientations pédagogiques principales, tandis que les élèves ont été interrogés sur le contexte de leur classe et les pratiques qui y sont d'usage. Les échelles PISA de climat de discipline et de soutien de l'enseignant en cours de mathématiques sont des indicateurs de la structure et du soutien dans le modèle triarchique de la qualité de l'instruction. Le cycle PISA 2006 a inclus plusieurs indicateurs sur les pratiques pédagogiques novatrices en cours de sciences, et le cycle PISA 2012 en a fait de même en mathématiques (voir la section « Possibilités d'apprentissage et qualité de l'instruction : évaluation de l'environnement d'apprentissage »). Des informations détaillées sur les convictions des enseignants, entre constructivisme et transmission directe par exemple, et sur leurs activités de développement professionnel viendront assurément affiner le trait. À l'occasion du cycle PISA 2012, certains pays devraient y parvenir en établissant un lien entre l'enquête PISA 2012 et l'enquête TALIS sur les enseignants en 2013. Ces données agrégées à l'échelle des établissements d'enseignement pourront donc alimenter des analyses secondaires des résultats de l'enquête PISA.

Intrants et processus à l'échelle des élèves

La performance des individus à l'âge de 15 ans n'est pas seulement le fruit des effets cumulés de la scolarité, elle est également le fruit des expériences d'apprentissage vécues dans le milieu familial. Les parents plus instruits sont capables d'offrir de meilleures possibilités d'apprentissage à leurs enfants. Ils sont aussi capables de leur donner un meilleur accès à l'écrit, au voyage et à d'autres ressources qui excitent la curiosité de leurs enfants. Selon certaines études, le fait que les parents nourrissent de grandes ambitions au sujet de la scolarité de leurs enfants et s'intéressent à leur travail scolaire est en corrélation avec leur réussite scolaire ; il en va de même avec la participation des parents aux conférences scolaires et leur implication dans les devoirs (Alexander *et al.*, 2007 ; Christenson, 2004 ; Hoover-Dempsey et Sandler, 1997 ; Ma, 1999 ; Sui-Chu et Willms, 1996 ; Wang *et al.*, 1993).

L'un des objectifs de la collecte de données sur le milieu familial des élèves est de tenir compte de ses impacts sur l'apprentissage lors de l'estimation des effets à l'échelle des établissements d'enseignement. Outre le questionnaire « Élève », le questionnaire « Parents » proposé à titre d'option internationale permet de recueillir des informations sur plusieurs variables pertinentes, notamment les attentes des parents concernant la carrière de leurs enfants, leur implication dans l'établissement d'enseignement et les matières scolaires évoquées en famille. Ensemble, ces instruments permettent de procéder à la triangulation du soutien des parents et de leurs attentes académiques sous divers angles. Les données sur le matériel informatique que les élèves ont à leur disposition, sur l'usage qu'ils en font – y compris à domicile – et sur leurs attitudes et compétences en la matière qui sont recueillies *via* le questionnaire sur la maîtrise des TIC contribueront en particulier à expliquer les résultats aux épreuves cognitives informatisées.

Outre le milieu socio-économique et familial des élèves, leur contexte linguistique semble pertinent. Grandir dans un milieu bilingue peut favoriser la sensibilisation linguistique et faciliter l'apprentissage de langues étrangères (Hesse *et al.*, 2008). Toutefois, la maîtrise imparfaite de la langue d'enseignement durant les premières années de scolarité peut



aussi avoir un impact négatif sur l'apprentissage dans d'autres matières et nuire au parcours scolaire des élèves à long terme (Schmid, 2001 ; Stanat et Christensen, 2006). D'une part, les migrants et les minorités ethniques peuvent avoir à s'adapter à un contexte culturel qui ne leur est pas familier à l'école (Berry, 1980, 1990 ; Hovey et King, 1996 ; Liebkind, 1996) et s'exposer à la discrimination (Amiot et Bourhis, 2005 ; Perreault et Bourhis, 1999). D'autre part, ces groupes se distinguent souvent par de plus hautes aspirations et une plus grande motivation que leurs condisciples qui ne vivent pas dans des minorités ethniques (Kao et Tienda, 1998 ; Krahn et Taylor, 2005 ; Stanat, 2006 ; Stanat et Christensen, 2006). Au sujet des élèves issus de l'immigration, l'enquête PISA cherche à identifier les différents intrants, processus et extrants de l'éducation pour isoler les effets de la migration des effets socio-économiques et linguistiques. Elle cherche aussi à isoler les effets individuels des effets collectifs ou contextuels. Le cycle PISA 2012 tente d'accroître le pouvoir explicatif de l'étude en incluant des indicateurs d'acculturation et de proximité culturelle perçue ainsi que des informations détaillées sur le contexte linguistique via le questionnaire sur le parcours scolaire. Le questionnaire facultatif « Parents » fournit des informations supplémentaires sur l'ascendance allochtone, et le questionnaire facultatif sur le parcours scolaire, des informations supplémentaires sur le contexte linguistique familial.

Les processus généraux de niveau Élève concernent essentiellement le temps d'apprentissage. Les possibilités formelles d'apprentissage sont d'ordinaire obligatoires, mais les élèves peuvent ne pas en tirer pleinement parti, que ce soit parce qu'ils sèchent les cours, qu'ils arrivent en retard ou qu'ils ne sont pas attentifs en classe. L'absentéisme et le temps consacré aux tâches influent fortement sur la performance des élèves et sont en corrélation avec les taux d'abandon, de criminalité et de toxicomanie (voir, par exemple, Baker *et al.*, 2001 ; Lee et Burkam, 2003 ; McCluskey *et al.*, 2004 ; Wilmers *et al.*, 2002). De plus, ces comportements sont également pertinents sous forme d'indicateur agrégé de l'environnement d'apprentissage en milieu scolaire. En revanche, les élèves peuvent allonger leur temps d'apprentissage et diversifier leurs possibilités d'apprentissage en se livrant à des activités d'apprentissage individuelles ou à des activités extrascolaires, par exemple les clubs académiques, les concours, le bénévolat et les clubs de discussion. Les élèves issus de milieux sociaux défavorisés semblent bénéficier particulièrement de la participation à ce type d'activités (voir, par exemple, McComb et Scott-Little, 2003). Dans certains pays, suivre des cours payants de soutien scolaire (comme ceux proposés dans les *juku* au Japon) est courant et constitue une partie importante du système d'éducation (Baker et LeTendre, 2005).

Les ambitions scolaires sont une variable prédictive probante du niveau de formation futur des élèves (voir, par exemple, Thiessen, 2007) ; elles varient entre les groupes ethniques en termes de qualité et de stabilité (voir, par exemple, Kao et Tienda, 1998 ; Mau et Heim Bikos, 2000). La motivation des élèves et leurs attitudes à l'égard de l'apprentissage sont des intrants importants de ce niveau, qui peuvent également être considérés comme des résultats non cognitifs. Comme les styles d'apprentissage des élèves, elles sont en grande partie spécifiques aux matières ; c'est la raison pour laquelle nous y reviendrons dans la section suivante « Les conditions d'apprentissage au service de la culture mathématique ». Des caractéristiques générales, indépendantes des matières, des élèves peuvent en revanche expliquer la performance des élèves, en particulier des compétences transversales telles que la résolution de problèmes. Le cycle PISA 2012 inclut un indicateur sur l'ouverture des élèves à l'apprentissage et à l'exploration, ainsi qu'une taxinomie des approches de résolution de problèmes et un indicateur sur les stratégies lors des tests.

Concernant l'élaboration des variables, contrairement aux caractéristiques spécifiques au milieu et au soutien familial qu'il est plus judicieux de dériver du questionnaire « Parents », la plupart des caractéristiques constituant le profil d'apprenant des élèves sont, comme par le passé, évaluées sur la base du questionnaire « Élève » (d'autant que le questionnaire « Parents » est une option internationale en 2012). Deux des options internationales proposées lors du cycle PISA 2012 fourniront des informations supplémentaires, en l'occurrence les questionnaires sur la maîtrise des TIC et sur le parcours scolaire. Les variables prédictives de niveau Élève les plus probantes de l'apprentissage (la capacité intellectuelle et les connaissances antérieures) ne sont toutefois pas faciles à évaluer dans l'enquête PISA. Toutefois, des extensions longitudinales de l'enquête PISA, qui ont été menées avec succès dans certains pays, ont réussi à utiliser les capacités intellectuelles et les connaissances antérieures pour estimer l'impact de variables scolaires. Il y a lieu d'envisager la possibilité d'opter pour des extensions longitudinales si les responsables politiques veulent être bien informés des effets des établissements d'enseignement et de la scolarité (voir les conclusions dans la section « Garantir la validité des instruments PISA, leur pouvoir explicatif et leur pertinence pour l'action publique »).

LES CONDITIONS D'APPRENTISSAGE AU SERVICE DE LA CULTURE MATHÉMATIQUE

Les mathématiques sont le domaine majeur d'évaluation du cycle PISA 2012, ce qui a des implications claires pour l'élaboration des variables contextuelles propres aux élèves et au système d'éducation. Conformément à la taxinomie des variables de l'éducation de niveau Élève présentée au tableau 6.1, l'enquête PISA cherche à décrire des aspects



importants de l'affect des élèves et leur expérience des mathématiques dans le cadre scolaire et en dehors (voir, par exemple, l'expérience d'approches différentes à l'égard de l'enseignement des mathématiques, les approches d'apprentissage préférées), leur motivation, et leur intérêt et leur engagement en la matière. Au niveau institutionnel (voir les variables de niveau Classe et Établissement dans le tableau 6.1), l'enquête PISA cherche à décrire des aspects importants de l'apprentissage et de l'instruction en mathématiques, et notamment à analyser la relation entre les stratégies d'apprentissage et d'enseignement et la performance des élèves, ainsi que la relation entre l'organisation des établissements d'enseignement, les structures et l'engagement actif des élèves dans l'apprentissage. Ces objectifs cadrent bien avec la « Stratégie de développement de l'enquête PISA à plus long terme », un texte dans lequel le Comité directeur PISA a affirmé dès 2005 qu'il était « concevable de faire de l'efficacité et de l'efficacé des processus éducatifs la thématique majeure du cycle PISA 2012 », car cette analyse « se prêterait particulièrement bien aux mathématiques, le domaine d'évaluation le plus fortement corrélé aux programmes de cours de l'enquête PISA »⁶.

Dans les sections suivantes sont décrits les fondements de l'analyse des diverses facettes de la culture mathématique, dont les conditions essentielles à l'échelle des élèves et des systèmes. De plus, les considérations conceptuelles sont accompagnées de recommandations concernant la conception des questionnaires contextuels du cycle PISA 2012.

La culture mathématique : Un défi en matière d'instruction et d'évaluation

Comme la culture mathématique est l'une des principales conditions de réussite dans la société moderne de la connaissance à forte composante technologique, c'est l'une des priorités de la plupart des systèmes d'éducation. Les compétences en mathématiques ont été et restent un domaine essentiel de la quasi-totalité des évaluations cognitives à grande échelle, qu'elles soient nationales ou internationales. Dans le monde entier, les responsables politiques se soucient énormément de la qualité de l'instruction des mathématiques. Toutefois, la signification même d'instruction de qualité a beaucoup changé ces 20 dernières années (voir Schoenfeld, 2006). Ce mouvement de réforme a de multiples origines, du pragmatisme de Dewey à la notion de développement cognitif du psychologue russe Vygotsky, en passant par la « *Reformpädagogik* » (réforme pédagogique) allemande, l'épistémologie constructiviste moderne et la théorie de l'autorégulation de l'apprentissage. À partir de ces sources, des concepts non mécaniques d'apprentissage et d'enseignement ont été inventés, et ont fini par s'infiltrer dans les pratiques pédagogiques des professeurs de mathématiques.

Le document « Curriculum and Evaluation Standards for School Mathematics », publié en 1989 aux États-Unis par une association d'enseignants, est sans doute le texte fondateur de la conception moderne de l'enseignement des mathématiques ; il y a été fait largement écho dans le monde entier. Ce document décrit « cinq objectifs fixés à tous les élèves : *i*) apprendre à valoriser les mathématiques ; *ii*) devenir confiants en leur capacité de faire des mathématiques ; *iii*) devenir des acteurs de la résolution de problèmes mathématiques ; *iv*) apprendre à communiquer de façon mathématique ; et *v*) apprendre à raisonner de façon mathématique » (NCTM, 1989, p. 5). Les normes du NCTM marquent un changement net dans la façon de concevoir l'enseignement des mathématiques, car les facultés générales de raisonnement mathématique viennent remplacer les objectifs d'apprentissage axés sur les contenus que l'on utilisait auparavant pour caractériser la finalité de l'éducation. L'enseignement des mathématiques vise désormais à aiguïser les facultés de raisonnement mathématique et à améliorer leur application dans des problèmes tirés du monde réel, bien au-delà de ce qui se faisait auparavant, à savoir inculquer des connaissances à restituer et des compétences procédurales. Des recherches plus récentes ont développé et étoffé ces compétences au travers de travaux théoriques et d'études empiriques (Niss, 2003 ; Blomhøj et Jensen, 2007). Plus récemment encore, des avancées similaires ont été faites via la *Common Core State Standards Initiative* (CCSSI) (2010) pour les mathématiques aux États-Unis. La version préliminaire du document de la CCSSI contient non seulement des normes de contenus mathématiques, mais aussi des pratiques mathématiques qui se basent explicitement sur les normes du NCTM évoquées ci-dessus.

Le cadre PISA d'évaluation des mathématiques s'inspire de ces avancées. Depuis 10 ans, l'enquête PISA a compté parmi les acteurs majeurs de la promotion de l'approche basée sur la modélisation et le raisonnement à l'égard des mathématiques, qui bénéficie du soutien sans faille au niveau mondial des responsables politiques et des experts spécialisés dans l'enseignement des mathématiques. Le cadre PISA 2012 d'évaluation des mathématiques (voir le chapitre 1) ne reflète pas les disciplines traditionnelles des mathématiques telles qu'elles sont enseignées (l'algèbre, la géométrie, les statistiques, etc.), mais accorde la priorité aux « idées majeures » qui guident la compréhension conceptuelle et s'attache aux compétences mathématiques, bien au-delà des savoirs et savoir-faire techniques.

Lors de la conceptualisation de l'enquête PISA, l'idée fondamentale était d'évaluer le raisonnement mathématique en contexte. Dans la version préliminaire du cadre d'évaluation des mathématiques du cycle PISA 2012, la culture mathématique se définit comme « l'aptitude d'un individu à formuler, employer et interpréter des mathématiques



dans un éventail de contextes, autrement dit sa capacité à se livrer à un raisonnement mathématique et à utiliser des concepts, procédures, faits et outils mathématiques pour décrire, expliquer et prévoir des phénomènes. Elle aide les individus à comprendre le rôle que les mathématiques jouent dans le monde et à se comporter en citoyens constructifs, engagés et réfléchis, c'est-à-dire capables de poser des jugements et de prendre des décisions en toute connaissance de cause » (voir le chapitre 1). Le concept n'est guère différent de celui qui a été élaboré, puis utilisé pour concevoir les épreuves PISA des cycles antérieurs, où les items étaient conçus pour évaluer « la mesure dans laquelle [les élèves] sont capables de mettre en œuvre leurs compétences mathématiques pour résoudre des problèmes de la vie courantes » (OCDE, 2004, p. 40). Le cadre du cycle PISA 2012 (voir le chapitre 1) base toujours l'évaluation des savoir-faire mathématiques des adolescents de 15 ans sur les phases du processus de modélisation mathématique et retient une série de compétences mathématiques fondamentales : *communication, représentation, élaboration de stratégies, mathématisation, raisonnement et argumentation, utilisation d'opérations et d'un langage symbolique, formel et technique, et utilisation d'outils mathématiques*. De plus, les items sont conçus en fonction de quatre grandes catégories de contenus mathématiques dans l'enseignement secondaire : *la quantité, l'incertitude et les données, les variations et les relations, et l'espace et les formes*. L'évaluation PISA de la culture mathématique continuera donc à moyen terme à aller bien au-delà des connaissances et compétences mathématiques purement techniques.

Les épreuves cognitives PISA couvrent donc trois des cinq objectifs de l'enseignement des mathématiques cités ci-dessus, à savoir la capacité des élèves à se livrer à un raisonnement mathématique, à résoudre des problèmes et à communiquer. Les deux autres objectifs, valoriser les mathématiques et avoir confiance en sa capacité à faire des mathématiques, sont évalués à titre de résultats non cognitifs dans le questionnaire « Élève », comme les stratégies en rapport avec les mathématiques. Ces *constructs* sont considérés comme des résultats de l'enseignement des mathématiques et aideront de surcroît à expliquer les écarts de performance dans les épreuves PISA de mathématiques. La section suivante en explique les fondements conceptuels et les indicateurs.

Indicateurs de résultats en rapport avec la culture mathématique : Stratégies, convictions et motivation

Stratégies et métacognition

Les professeurs de mathématiques veulent non seulement connaître la nature et le nombre des problèmes que les élèves sont capables de résoudre, mais également leur *cheminement* pour résoudre ces problèmes (les stratégies qu'ils appliquent pour étudier leurs leçons de mathématiques, se préparer à des tests ou encore aborder des problèmes). Les stratégies d'apprentissage des mathématiques et de résolution de problèmes sont des résultats importants de l'enseignement et une condition essentielle à la poursuite de l'apprentissage en mathématiques – au même titre que les stratégies d'apprentissage sur la base de textes en compréhension de l'écrit. Les élèves qui connaissent et maîtrisent des stratégies sont capables d'utiliser leur culture mathématique dans de nouveaux contextes et de nouvelles tâches.

En compréhension de l'écrit, l'enquête PISA a utilisé les autoévaluations des élèves au sujet de trois stratégies d'apprentissage, à savoir la mémorisation (apprendre des mots clés, réciter des contenus), l'élaboration (établir des liens entre des domaines connexes, réfléchir à d'autres solutions) et le contrôle (des stratégies métacognitives impliquant une planification, un suivi et une régulation). Ces échelles ont été adaptées aux mathématiques lors du cycle PISA 2003. L'échelle d'élaboration a été constituée sur la base d'items adaptés de l'évaluation de la compréhension de l'écrit, tels que « Quand j'apprends des mathématiques, j'essaie d'établir des relations entre ce travail et des choses apprises dans d'autres matières », ainsi que d'items spécifiques tels que « Quand je résous des problèmes de mathématiques, j'imagine souvent de nouvelles façons de trouver la réponse » et « Quand je résous un problème de mathématiques, je réfléchis souvent à la façon dont on pourrait appliquer la solution à d'autres cas intéressants ». Cette nouvelle version de l'échelle s'est révélée valide : c'est une variable prédictive probante de la performance en mathématiques dans les pays participants. En revanche, l'échelle de mémorisation n'a pas été retenue à titre d'échelle distincte valide, et des résultats mitigés ont été obtenus pour les stratégies de contrôle que les élèves ont été invités à évaluer (OCDE, 2005b, p. 297 ; Vieluf *et al.*, 2009a, 2009b).

C'est à dessein que les items relatifs aux stratégies d'élaboration en mathématiques se situent dans le contexte de la résolution de problèmes. Des études comparatives internationales montrent en effet que l'enseignement des mathématiques est essentiellement défini en termes de résolution de problèmes dans le monde entier (Christiansen et Walther, 1986 ; Hiebert *et al.*, 2003 ; Stigler et Hiebert, 1999). Lorsque les élèves travaillent seuls, en petits groupes ou tous ensemble, ils doivent souvent résoudre des problèmes ou effectuer des tâches de routine. Dans de nombreux pays, les enseignants se servent d'exemples de problèmes, même lorsqu'ils présentent un nouveau contenu, pour que leurs élèves découvrent de nouvelles méthodes de résolution de problèmes. Étudier ou apprendre en mathématiques signifie



donc essentiellement résoudre des problèmes. Ce sont les stratégies de résolution de problèmes qui sont au cœur de la culture mathématique, et non les stratégies d'apprentissage comme en compréhension de l'écrit. Des études sur les aspects cognitifs de l'apprentissage des mathématiques les ont redéfinies en tant qu'instances de ce que l'on appelle couramment la métacognition (Desoete et Veenman, 2006 ; Garofalo et Lester, 1985 ; Schneider et Artelt, à paraître ; Schoenfeld, 1992). En général, la métacognition désigne le fait que l'individu connaît ses stratégies et processus cognitifs et y réfléchit, comme le définit Flavell dans un texte désormais de référence (1979).

Concernant l'élaboration des échelles, l'échelle de stratégies d'élaboration créée à l'occasion du cycle PISA 2003 (autoévaluation des stratégies d'apprentissage en mathématiques et de résolution de problèmes) a été conservée pour permettre des analyses de tendances. Toutefois, comme Schneider et Artelt (à paraître) le soulignent, les indicateurs de métacognition dérivés des évaluations que font les élèves de leurs stratégies et de leur usage ne sont guère probants. De plus, il est possible d'envisager à l'avenir l'utilisation d'indicateurs dérivés des fichiers journaux des épreuves cognitives informatisées ou d'indicateurs dérivés des connaissances métacognitives déclarées comme ceux élaborés sur la base de l'évaluation de la métacognition en compréhension de l'écrit lors du cycle PISA 2009.

Motivation et intentions

À l'heure où les formations en sciences, en technologie, en ingénierie et en mathématiques attirent peu les étudiants, surtout de sexe féminin – un problème qui se pose dans de nombreux pays –, accroître l'intérêt pour les mathématiques et la motivation à l'idée d'apprendre en mathématiques, et faire naître des attitudes générales plus positives à l'égard des mathématiques, sont des priorités pour les responsables politiques. De plus, les éléments à l'appui d'une relation positive entre des attitudes favorables à l'égard des mathématiques et la performance en mathématiques abondent (voir, par exemple, la méta-analyse de Ma et Kishor, 1997).

Les théories sur la motivation des élèves font généralement la distinction entre la motivation intrinsèque et la motivation extrinsèque. La motivation intrinsèque provient de la gratification inhérente à une tâche ou à une activité. Lors du cycle PISA 2003, ce type de motivation a fait l'objet d'une échelle d'intérêt et de plaisir des mathématiques (sur la base d'items tels que « Je m'intéresse aux choses que j'apprends en mathématiques »). La littérature suggère que la motivation intrinsèque a un impact sur l'engagement des élèves et le temps qu'ils consacrent aux tâches, leurs activités d'apprentissage, leur performance et leur choix de carrière, et qu'elle peut être forgée en classe (Kunter, 2005 ; Rakoczy *et al.*, 2008 ; Ryan et Deci, 2000). Par contraste, la motivation extrinsèque provient de la gratification attendue. Il est établi qu'elle est une variable prédictive probante du choix de cours et de carrière, et de la performance (Wigfield *et al.*, 1998). Lors du cycle PISA 2003, une échelle de motivation instrumentale a été élaborée au sujet de l'apprentissage des mathématiques, à partir d'items tels que « Pour moi, cela vaut la peine d'apprendre les mathématiques, car cela améliore mes <perspectives, chances> de carrière professionnelle ».

Des modèles plus sophistiqués de la maîtrise comportementale font la distinction entre des dispositions plus générales telles que la motivation intrinsèque et extrinsèque et les intentions. Les intentions sont plus proches du processus réel de prise de décision et sont plus pertinentes pour déterminer dans quelle mesure les élèves utilisent les mathématiques dans leur vie et les intègrent dans leurs projets d'avenir. Ces items sont vraisemblablement moins biaisés en raison de réponses spécifiques aux cultures ou de normes liées à des groupes de pairs. Les intentions à court terme peuvent être évaluées sur la base d'une échelle dérivée d'items tels que « Je vais travailler dur pour bien apprendre en mathématiques » (Lipnevich *et al.*, à paraître). Les intentions à long terme ou les orientations à l'avenir peuvent être rapportées sur une version adaptée de l'échelle élaborée au sujet des sciences lors du cycle PISA 2006 sur la base d'items tels que « J'aimerais exercer une profession dans laquelle interviennent les mathématiques ». Cette échelle permettrait de déterminer dans quelle mesure les pays peuvent espérer amener des sujets très performants à opter pour des professions en rapport avec les mathématiques, et d'étudier les inégalités d'accès dans les professions en rapport avec les sciences, la technologie, l'ingénierie et les mathématiques.

Concernant l'élaboration des variables, ces quatre échelles, à savoir l'intérêt et le plaisir des mathématiques, la motivation instrumentale relative à l'apprentissage en mathématiques, les intentions à court terme et les intentions à long terme, suffiront pour répondre à des questions sur la motivation des élèves en mathématiques qui sont pertinentes pour l'action publique. L'inclusion de six échelles d'attitudes dans le questionnaire « Élève » et l'intégration de plusieurs items d'attitude dans les épreuves cognitives, comme lors du cycle PISA 2006, surchargeraient inutilement les élèves. Toutefois, des approches novatrices pourraient être adoptées à titre expérimental pour réduire les biais culturels et garantir la comparabilité des résultats entre les cultures (voir la section « Extension de l'échantillon pour étudier la variation au niveau Établissement [option internationale] »).



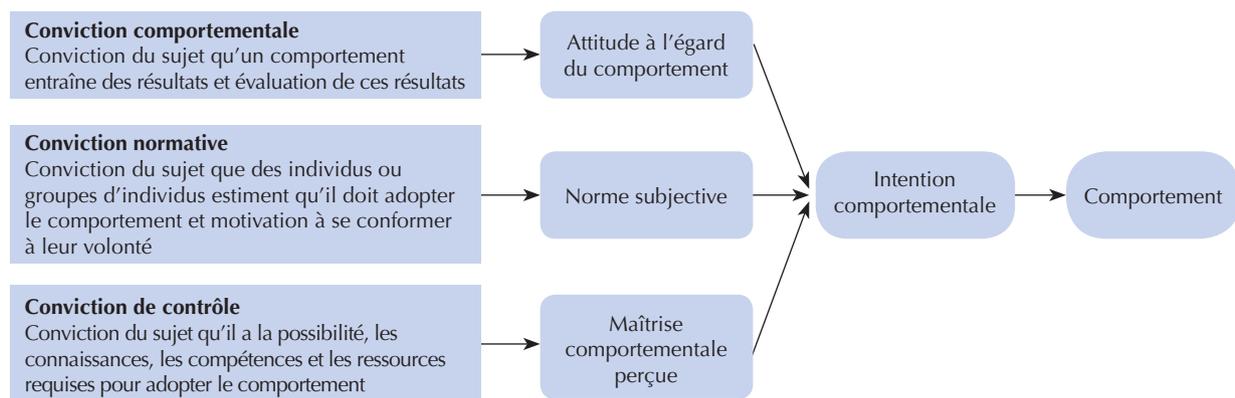
Image de soi et comportement prévu

La façon dont les élèves se considèrent est une variable prédictive probante de leurs actes et de leurs décisions face à des tâches et des situations difficiles (Bandura, 1977). L'efficacité perçue se définit comme la perception d'un individu de sa capacité à mener à bien certaines tâches. Quant à la perception de soi, elle renvoie plus largement à la façon dont les individus évaluent leurs attributs. Selon des études sur l'enseignement des mathématiques, l'efficacité perçue en mathématiques est une variable prédictive probante des notes en mathématiques, de la performance en résolution de problèmes, de l'intérêt pour les mathématiques et – une fois encore – des choix de carrière (Britner et Pajares, 2006 ; Pajares et Miller 1997 ; Turner *et al.*, 2004). Lors du cycle PISA 2003, deux indicateurs ont été élaborés : la perception de soi en mathématiques (sur la base d'items tels que « J'apprends vite en mathématiques ») et la perception des capacités personnelles à surmonter les difficultés en mathématiques (voir Lee, 2009 ; OCDE, 2005b). Les élèves ont plus précisément été invités à indiquer dans quelle mesure ils se sentent sûrs d'eux à l'idée d'effectuer huit tâches de mathématiques dont la difficulté varie : certaines se basent sur des procédures techniques directes (résoudre une équation du type : $3x + 5 = 17$) et d'autres requièrent une modélisation mathématique (calculer la consommation d'essence d'une voiture) (OCDE, 2005b, p. 292). La perception de soi en mathématiques et la perception des capacités personnelles à surmonter des difficultés en mathématiques se sont toutes deux révélées équivalentes entre les cultures et il est apparu qu'elles étaient des variables prédictives probantes de la performance des élèves (Vieluf *et al.*, 2009a, 2009b).

La mesure dans laquelle les élèves s'estiment confiants dans leur capacité à mener à bien des tâches mathématiques et la mesure dans laquelle ils valorisent les mathématiques sont des variables très pertinentes pour prévoir et expliquer le comportement des élèves en mathématiques, en cours et dans leur choix de carrière. Un certain nombre de modèles Attentes-Valeur ont été proposés, tant en psychologie qu'en économie, pour intégrer les deux aspects de la prise de décision. L'un de ces modèles est celui de la théorie du comportement planifié d'Ajzen (1991), selon laquelle le comportement volitif est déterminé par des attitudes spécifiques et des normes subjectives (côté valeur), et par le contrôle comportemental perçu (côté attentes). (Pour une application aux mathématiques dans les recherches interculturelles, voir Lipnevich *et al.*, à paraître.) Selon Ajzen (1991), la manipulation de ces variables prédictives permet d'accroître la probabilité qu'un individu ait l'intention de poser un acte, ce qui entraîne l'augmentation de la probabilité qu'il pose effectivement cet acte (voir la figure 6.1).

■ Figure 6.1 ■

La théorie du comportement planifié



Dans ce contexte, les attitudes, la maîtrise comportementale perçue et les normes subjectives peuvent être des variables prédictives de la volonté des élèves de consacrer du temps à leurs devoirs de mathématiques, de poser des questions en classe ou de se livrer à des exercices de relaxation pour lutter contre leur anxiété et, donc, d'améliorer leur performance en mathématiques.

Concernant l'élaboration des variables, l'application du modèle complet d'Ajzen accroît sensiblement le pouvoir analytique de l'enquête pour expliquer les efforts des élèves en mathématiques, leur comportement en mathématiques et, indirectement, leurs résultats, si possible. Toutefois, en raison des coûts et des ressources à mobiliser pour évaluer toutes les facettes du modèle, cette extension n'est soumise qu'à des sous-groupes d'élèves, moyennant la rotation des items (voir la section « Rotation des items dans les extensions du questionnaire "Élève" et les instruments informatisés »).



Dans le questionnaire « Élève », les deux indicateurs d'image de soi en mathématiques sont utilisés et l'indicateur d'anxiété en mathématiques est conservé. Toutefois, de nouveaux items ont été expérimentés en raison de doutes sur la validité interculturelle et le pouvoir explicatif incrémentiel (Lee, 2009 ; Vieluf *et al.*, 2009a, 2009b).

Possibilités d'apprentissage et qualité de l'instruction : Évaluation de l'environnement d'apprentissage

L'enquête PISA cherche à décrire les conditions d'apprentissage dans le domaine majeur d'évaluation du cycle pour deux raisons. En premier lieu, l'enquête PISA contribue à décrire et comparer les environnements d'apprentissage des élèves âgés de 15 ans dans les différents pays, sur la base de variables dont la littérature établit qu'elles sont des indicateurs probants de la qualité de l'instruction. En deuxième lieu, ces variables contribuent à expliquer les résultats des élèves – pas dans tous les pays, mais dans ceux où la plupart des élèves fréquentent un établissement échantillonné à deux reprises au moins – ce qui est le cas dans deux tiers des pays qui ont participé au cycle PISA 2006.

Les environnements d'apprentissage qui aident les élèves à développer leur culture mathématique – telle qu'elle se conçoit dans l'enquête PISA – sont décrits dans la littérature comme des cadres d'apprentissage en contexte (Schoenfeld, 2006), où l'enseignement est basé sur la discussion (Boaler et Greeno, 2000 ; Stein *et al.*, 2008) et qui incluent la modélisation mathématique (Blum et Leiss, 2005) à titre d'activité essentielle. Dans ses travaux précurseurs, Schoenfeld (1987, 1992) a plaidé en faveur d'une culture mathématique qui permette un apprentissage éclairé, qui établisse des liens entre des concepts mathématiques et qui rapporte ces liens à la vie courante des élèves – par comparaison avec l'instruction classique qui vise la maîtrise de formules, d'algorithmes, de définitions et autres contenus techniques. La « conception moderne de l'enseignement efficace des mathématiques » (Stein *et al.*, 2008) consiste notamment « à présenter aux élèves des problèmes mathématiques complexes plus réalistes, à les amener à se renvoyer la balle pour travailler sur ces problèmes, puis à partager les stratégies et les solutions des uns et des autres lors d'une discussion orchestrée par l'enseignant avec toute la classe » (p. 315).

Comme l'enquête PISA entend couvrir les aspects majeurs de l'environnement d'apprentissage des mathématiques dans ses questionnaires, cette conception moderne de l'enseignement des mathématiques est un défi. Les items sur les contenus ne suffisent pas pour décrire les processus d'enseignement et d'apprentissage propices au développement de la culture mathématique. Tout d'abord, parce que la culture mathématique telle qu'elle se conçoit dans l'enquête PISA renvoie à un rendement cumulé, qui dépend plus de compétences et de processus mathématiques fondamentaux et, donc, moins de contenus spécifiques qui auraient été enseignés. Ensuite, parce que la qualité de l'enseignement des mathématiques dépend moins des contenus que de la clarté et de la structure des cours, du soutien aux élèves et des difficultés cognitives (voir la section « Intrants et processus à l'échelle des classes »). C'est pourquoi l'enquête PISA a redéfini la notion de possibilité d'apprentissage, un *construct* important qui caractérise les environnements d'apprentissage dans d'autres études internationales sur les acquis des élèves. De plus, les données sur les possibilités d'apprentissage obtenues auprès des élèves peuvent être étudiées à l'échelle des établissements d'enseignement, grâce à la conception des instruments PISA.

Possibilités d'apprentissage : couverture des catégories de contenus et des types de problèmes

La notion de possibilité d'apprentissage a été introduite par John Carroll au début des années 60, initialement pour indiquer si les élèves avaient eu suffisamment de temps et reçu une instruction adéquate pour apprendre (Carroll, 1963 ; Abedi *et al.*, 2006). Elle est depuis lors devenue un concept important dans les évaluations internationales des acquis des élèves (Husén, 1967 ; Schmidt et McKnight, 1995 ; Schmidt *et al.*, 2001), et il est établi qu'elle est en forte corrélation avec la performance des élèves, en particulier dans les comparaisons entre les pays (Schmidt et Maier, 2009, pp. 552-556). Parallèlement, la portée du *construct* s'est considérablement développée. Stevens (1993, pp. 233-234) a identifié les quatre types de variables les plus souvent cités dans la littérature au sujet des possibilités d'apprentissage :

- La couverture du contenu : ces variables permettent de déterminer si les élèves voient ou non le programme prévu dans une année d'études ou une matière.
- L'exposition au contenu : ces variables portent sur le temps prévu et consacré à l'instruction, et sur le caractère approfondi de l'enseignement dispensé.
- L'importance du contenu : ces variables montrent l'importance accordée aux thématiques du programme et identifient les élèves sélectionnés pour recevoir une instruction privilégiant les compétences d'ordre inférieur (mémorisation par cœur) ou d'ordre supérieur (résolution de problèmes).
- La qualité de l'instruction : ces variables montrent en quoi les pratiques pédagogiques en classe (la présentation des cours) influent sur la performance académique des élèves.



Pour certains auteurs, les possibilités d'apprentissage sont devenues plus ou moins synonymes de qualité de l'instruction. Schmidt et Maier (2009) précisent toutefois dans leur analyse que les possibilités d'apprentissage sont un concept relativement simple : « Ce que les élèves apprennent à l'école dépend de ce qu'on leur enseigne » (p. 541), et ils se concentrent à dessein sur les possibilités d'apprentissage « au sens le plus strict, c'est-à-dire l'exposition des élèves au contenu » (p. 542).

Schmidt et Maier reconnaissent que les possibilités d'apprentissage sont un *construct* direct, mais qu'elles sont difficiles à mesurer. Pour expliquer les différences de matières vues, il est d'usage de demander aux enseignants et/ou aux élèves si des matières du programme prévu ont été enseignées ou non, parfois avec l'appui des journaux de classe (Rowan *et al.*, 2004). De plus, il est demandé à des experts spécialisés dans les programmes de cours si certains éléments des programmes sont couverts dans des documents tels que les syllabus, les manuels, etc. (le programme prévu) et, dans l'affirmative, de quelle manière. Divers indicateurs ont été dérivés de ces données brutes. Dans de nombreux cas, les matières enseignées sont jugées à deux égards, en termes de thématique et de niveau, alors qu'à l'échelle des systèmes, ce sont des indicateurs de cohérence, de rigueur et de priorité qui ont été dérivés (Schmidt et Maier, 2009).

Concernant l'élaboration des variables, l'évaluation PISA des possibilités d'apprentissage s'écarte des approches adoptées dans d'autres études, car la culture mathématique n'est pas mesurée en termes de contenus, mais sur la base de sept facultés mathématiques fondamentales et de quatre catégories de contenu. Dans l'enquête PISA, l'opérationnalisation des variables relatives aux possibilités d'apprentissage repose essentiellement sur les appréciations des élèves.

Le cycle PISA 2012 vise à déterminer le niveau de possibilités d'apprentissage dans les pays (voire dans les établissements d'enseignement). Les élèves se voient présenter des tâches mathématiques élaborées avec soin – certaines faisant intervenir les facultés mathématiques et les catégories de contenu décrites dans le cadre PISA de mathématiques, d'autres ressemblant à des tâches plus traditionnelles requérant la reproduction de procédures et la restitution de connaissances. Après avoir répondu à ces items, les élèves sont invités à indiquer s'ils ont déjà abordé de telles tâches au cours de mathématiques ou lors d'évaluations précédentes et, dans l'affirmative, à quelle fréquence. Il est donc possible d'évaluer les possibilités d'apprentissage – à l'échelle des pays, voire des établissements d'enseignement – d'une façon qui permet de différencier les types de problèmes et de contenus.

De plus, les élèves sont invités à indiquer dans quelle mesure certains concepts mathématiques leur sont familiers. Cet indicateur, ainsi que l'échelle d'efficacité perçue qui est décrite dans la section « Indicateurs de résultats en rapport avec la culture mathématique : stratégies, convictions et motivation », peuvent être utilisés comme indicateurs proximaux des possibilités d'apprentissage.

Temps d'apprentissage

Après Carroll (1963), de nombreuses études ont montré que le temps d'apprentissage était une variable prédictive majeure des résultats des élèves dans de nombreuses matières, y compris en mathématiques. Le cycle PISA 2012 a utilisé un indicateur de 2003 pour permettre la réalisation d'analyses tendancielles. Les élèves ont indiqué le nombre moyen d'heures qu'ils consacrent par semaine à leurs devoirs, à des cours de soutien ou de perfectionnement, à des cours avec un professeur particulier, et à des cours en dehors de l'école. De plus, la durée et le nombre de cours de mathématiques par semaine ont été évalués *via* les questionnaires « Élève » et « Établissement ».

Qualité de l'instruction

Comme indiqué dans la section « Intrants et processus à l'échelle des classes », les études sur les cours de mathématiques dans l'enseignement secondaire suggèrent que les dimensions fondamentales de la qualité de l'instruction sont : *i*) la structure et la gestion efficace de la classe ; *ii*) le soutien des enseignants et l'orientation des élèves ; et *iii*) la difficulté et l'activation cognitive. Deux de ces dimensions sont couvertes par des échelles élaborées lors du cycle PISA 2003, à savoir le *climat de discipline* et le *soutien par l'enseignant* en cours de mathématiques, à partir du questionnaire « Élève ».

Le rapport international sur le cycle PISA 2003 (OCDE, 2004) montre que le *construct* du *climat de discipline* en cours de mathématiques est en forte corrélation avec la culture mathématique, alors que d'autres variables – telles que la taille des classes, les activités proposées en mathématiques dans l'établissement d'enseignement et le fait de ne pas pratiquer le regroupement par aptitudes – n'ont pas d'effet sensible après contrôle du milieu socio-économique. Ces constats rejoignent les conclusions d'études antérieures et la théorie exposée dans la section « Intrants et processus à l'échelle des classes », selon laquelle une bonne structure et une gestion de classe efficace sont des conditions essentielles à l'apprentissage des élèves. Toutefois, lors du cycle PISA 2003, la relation entre l'environnement d'apprentissage et des résultats non cognitifs tels que l'intérêt des élèves pour les mathématiques ou le plaisir que leur procurent les



mathématiques n'a pas été analysée. Depuis lors, Vieluf *et al.* (2009a) ont établi l'existence d'une relation positive entre le soutien des enseignants et l'intérêt des élèves pour les mathématiques, même après contrôle du milieu socio-économique.

La difficulté et l'activation cognitive sont des dimensions cruciales pour le développement de la culture mathématique, mais elles sont très difficiles à évaluer. Des concepts similaires, qui mettent en évidence l'importance de l'orchestration de l'enseignement des mathématiques d'une façon qui offre aux élèves de nombreuses possibilités de développer leurs compétences, sont évoqués sous un angle général dans la littérature sur l'enseignement des mathématiques (voir, par exemple, Blum et Leiss, 2007). Toutefois, aucun des indicateurs utilisés lors du cycle PISA 2003 ne permet de refléter cette dimension. Plusieurs approches ont été employées avec succès dans des études nationales et ont été expérimentées lors de l'essai de terrain du cycle PISA 2012 :

- Comme les devoirs sont une composante quasi universelle de l'enseignement des mathématiques, le traitement que les enseignants leur réservent en classe peut être utilisé comme indicateur de difficulté. Rakoczy *et al.* (2005) ont élaboré une échelle qui rend compte des processus relatifs aux devoirs sur la base d'items tels que « Le professeur de mathématiques tient à savoir comment nous avons résolu les problèmes de notre devoir » et « Lorsque nous en venons aux résultats des devoirs, nous tentons de comprendre les erreurs et de les corriger » (remarque : cette échelle ne porte pas sur les devoirs en tant que tels, ni sur le temps que les élèves y consacrent, mais sur la qualité du discours en classe à propos des résultats des devoirs).
- Les élèves peuvent aussi être interrogés sur le type d'items qu'ils rencontrent souvent en cours de mathématiques, par exemple : « Nous devons généralement réfléchir un moment pour résoudre les problèmes que nous donne le professeur de mathématiques » (échelle de tâches d'activation cognitive utilisée en Allemagne à l'occasion du cycle PISA 2003, Baumert *et al.*, 2008).

Pratiques pédagogiques et activités mathématiques des élèves

Une version à orientation comportementale du modèle triarchique de la qualité de l'instruction a été adoptée dans l'enquête TALIS de l'OCDE : les enseignants ont été invités à indiquer à quelle fréquence ils utilisent 13 pratiques pédagogiques différentes :

- Des pratiques structurantes (cinq items), par exemple annoncer clairement les objectifs d'apprentissage, résumer les cours précédents, passer les devoirs en revue, contrôler le cahier d'exercice et poser des questions aux élèves pour vérifier qu'ils comprennent la leçon.
- Des pratiques axées sur les élèves (quatre items), par exemple faire travailler les élèves en petits groupes pour qu'ils trouvent ensemble une solution à un problème ou à un exercice, pratiquer le regroupement par aptitudes, amener les élèves à s'autoévaluer et à participer à la planification des cours.
- Des activités de renforcement (quatre items), par exemple faire travailler les élèves à des projets d'une semaine au moins et leur demander de concevoir un produit, d'écrire une rédaction et de se livrer à des débats contradictoires.

Il ressort de l'analyse des résultats de la campagne définitive de l'enquête TALIS dans 23 pays que : *i*) les trois dimensions peuvent être différenciées entre les pays (c'est-à-dire que le modèle triarchique présente une certaine validité culturelle) ; *ii*) l'hypothèse selon laquelle les pratiques structurantes seraient associées à une meilleure discipline en classe (telle qu'elle est perçue par les enseignants) se confirme ; et *iii*) la participation à des activités de développement professionnel et la capacité d'enseigner dans des classes très performantes entraînent une augmentation de la fréquence de ces pratiques. Les professeurs de mathématiques et de sciences déclarent utiliser moins que leurs collègues qui enseignent d'autres matières les pratiques axées sur les élèves et les activités de renforcement (OCDE, 2009b).

Concernant l'élaboration des variables, les échelles TALIS de pratiques pédagogiques ont été adaptées pour pouvoir être utilisées lors du cycle PISA 2012, moyennant un bref ajout au questionnaire « Enseignant » de l'enquête TALIS, les items relatifs aux activités renforcées étant placés dans le contexte des mathématiques, par exemple utiliser les mathématiques pour résoudre un problème de la vie courante, établir un lien entre les mathématiques et d'autres matières, découvrir des règles à partir de structures mathématiques, démontrer un théorème mathématique, discuter d'exemples et de contre-exemples, utiliser de nombreuses représentations graphiques dans un problème mathématique, et comparer diverses méthodes de résolution ou plusieurs solutions pour la même tâche.

Seuls certains pays PISA peuvent établir un lien avec l'enquête TALIS, mais tous les élèves peuvent être invités à indiquer à quelle fréquence ils participent à de telles activités.



Soutien à l'enseignement et à l'apprentissage des mathématiques à l'échelle du système et des établissements d'enseignement

La plupart des variables de niveau Système et Établissement décrites dans les sections « Intrants et processus à l'échelle des systèmes » et « Intrants et processus à l'échelle des établissements d'enseignement » sont susceptibles d'être en corrélation avec la performance des élèves en mathématiques et/ou avec leurs attitudes à l'égard des mathématiques. Les indicateurs d'intrants et de processus spécifiques aux mathématiques peuvent être obtenus *via* l'agrégation des réponses des élèves aux items sur les possibilités d'apprentissage et la qualité de l'instruction (voir la section « Possibilités d'apprentissage et qualité de l'instruction : évaluation de l'environnement d'apprentissage ») à des niveaux supérieurs.

Toutefois, un certain nombre de variables spécifiques aux mathématiques à l'échelle des établissements d'enseignement ou du système façonnent directement les conditions d'enseignement et d'apprentissage des mathématiques. C'est pourquoi l'accent pourrait être mis sur les aspects suivants.

Réforme à l'échelle du système :

- initiatives, activités en vue d'améliorer le cadre scolaire et programmes de développement professionnel axés sur les mathématiques ;
- rôle des mathématiques dans l'admission des élèves, leur différenciation, leur évaluation et leur passage de classe, et l'évaluation des établissements d'enseignement ; et
- temps d'instruction prévu par an en mathématiques.

Réforme à l'échelle des établissements d'enseignement :

- offre de cours en mathématiques (cours de soutien et de perfectionnement, tutorat) ;
- qualification des professeurs de mathématiques ; et
- collaboration entre les professeurs de mathématiques.

Concernant l'élaboration des variables, les données peuvent être agrégées du niveau Établissement au niveau Système, car les chefs d'établissement sont interrogés sur toutes ces thématiques dans l'enquête PISA.

CONCEPTION DES QUESTIONNAIRES DU CYCLE PISA 2012

Répartition des contenus entre les cycles PISA : Une structure globale pour les variables générales, les variables spécifiques au domaine d'évaluation et les variables d'extensions thématiques

Parmi les facteurs pertinents pour comprendre la performance des élèves, leurs attitudes et leur comportement ainsi que le fonctionnement des systèmes d'éducation, certains sont directs (par exemple, les variables démographiques, le parcours scolaire antérieur, le temps d'enseignement et la taille des classes) et d'autres créés lors de précédents cycles PISA sont bien établis (le milieu socio-économique des élèves et leurs stratégies cognitives, ou encore le degré de participation des établissements d'enseignement à la prise de décision). Il est apparu que d'autres facteurs étaient moins faciles à cerner au travers de l'enquête PISA (par exemple, les politiques de responsabilisation à l'échelle des systèmes, les variables spécifiques aux enseignants, certains aspects de l'environnement d'apprentissage en classe ou les activités extrascolaires). Faire un choix parmi les nombreuses variables susceptibles d'être incluses est un processus complexe qui s'effectue sur la base des priorités fixées par les pays, ainsi que compte tenu de la littérature sur la pédagogie, comme indiqué dans les sections précédentes.

Dans la mesure où l'enquête PISA poursuit de nombreux objectifs et s'intéresse à de nombreux résultats de l'apprentissage des élèves, et que des informations sur les tendances sont nécessaires, les questionnaires contextuels doivent couvrir :

i) Des variables générales (lors de tous les cycles)

- Variables générales d'intrants :
 - Intrants de niveau Élève (année d'études, sexe et milieu socio-économique = niveau de formation et profession des parents, richesse, patrimoine culturel et ressources pédagogiques de la famille, données sur l'ascendance = ascendance autochtone ou allochtone, langue maternelle et âge d'arrivée dans le pays, soutien familial).
 - Contextes et intrants de niveau Établissement (taille de la communauté, ressources et qualifications du personnel enseignant).



- Variables générales de processus :
 - Processus de niveau Établissement (prise de décision, politiques d'admission, politiques d'évaluation, développement professionnel, engagement et moral des enseignants, relations entre enseignants et élèves, implication des parents).
 - Processus d'instruction (temps d'apprentissage, climat de discipline, soutien des enseignants).
- Variables générales de résultats :
 - Résultats non cognitifs d'ordre général – Engagement à l'égard de l'apprentissage (comportement : absentéisme ; objectif personnel : attentes en matière d'éducation ; motivation : engagement à l'égard de l'apprentissage ; affect : sentiment d'appartenance).

ii) Des variables tendancielle spécifiques au domaine d'évaluation (uniquement le domaine majeur d'évaluation, soit tous les neuf ans)

- Variables de résultats non cognitifs spécifiques au domaine d'évaluation (stratégies et métacognition, convictions, image de soi et motivation).
- Variables de processus spécifiques au domaine d'évaluation (possibilités d'apprentissage, qualité de l'instruction, soutien à l'échelle de l'établissement d'enseignement et du système).

iii) Des variables d'extensions thématiques (extensions lors de cycles précis) :

- Options internationales (questionnaires sur le parcours scolaire et la maîtrise des TIC, et questionnaire « Parents » en 2012, par exemple).
- Variables contextuelles dans d'autres domaines (les expériences en matière de TIC pertinentes pour les épreuves informatisées de résolution de problèmes, par exemple).
- Variables descriptives et explicatives pertinentes pour des rapports spécifiques (les motivations et les intentions en rapport avec les mathématiques en 2012, par exemple).
- Variables malléables à l'échelle des établissements d'enseignement (les politiques de différenciation et la certification des enseignants, par exemple), spécialement choisies à des fins de description ou d'inférence causale.

iv) Des données systémiques, recueillies *via* le projet INES ou un questionnaire « Système d'éducation »

- Résultats des établissements d'enseignement (délivrance des diplômes, par exemple).
- Ressources financières et humaines investies dans l'éducation.
- Accessibilité de l'éducation et scolarisation.
- Environnement d'apprentissage et organisation des établissements d'enseignement.

Lors de précédents cycles PISA – en particulier en 2006 –, les questionnaires ont largement négligé des *constructs* sans rapport avec le domaine majeur d'évaluation. La série de variables générales *i)* s'est parfois limitée à des intrants de niveau Élève. Le questionnaire « Élève » et, dans une certaine mesure, le questionnaire « Établissement », se sont concentrés sur des processus et des résultats non cognitifs spécifiques au domaine d'évaluation. Cette approche n'est pas appropriée lorsqu'il s'agit d'évaluer des tendances.

Trouver le juste équilibre entre les variables *i)*, *ii)*, *iii)* et *iv)* est essentiel pour la structure globale des questionnaires PISA et la réussite à long terme de l'enquête PISA.

Pour rendre compte de façon valide et fiable de tendances à l'échelle des pays, il est important de garder une série constante de variables de cycle en cycle, à titre de variables majeures de compte rendu et de bases d'estimation du niveau de compétence. Ces variables de contexte et d'intrant ne doivent donc pas changer.

Les responsables politiques n'ont pas seulement besoin d'informations tendancielle sur la performance des élèves, car les systèmes d'éducation poursuivent d'autres objectifs. Ils requièrent des informations sur des résultats non cognitifs, comportementaux ou sans rapport avec le domaine d'évaluation, à savoir les aspirations professionnelles (la profession que les élèves prévoient d'exercer à l'âge de 30 ans), l'engagement à l'égard de l'établissement d'enseignement (le sentiment d'appartenance), l'absentéisme et la motivation pour l'apprentissage (les efforts et la persévérance) – des résultats qui ont tous été couverts dans de précédents cycles PISA. Jusqu'ici, les auteurs de l'enquête PISA n'étaient pas favorables à leur utilisation, car ils craignaient que leur comparabilité entre cultures soit insuffisante. Toutefois, comme les informations sur les tendances sont disponibles pour un nombre croissant de cycles, la priorité va maintenant à l'évolution des taux



au sein des pays, plutôt qu'à la comparaison transversale des niveaux. Par exemple, l'éventuelle variation à la hausse ou à la baisse de l'engagement à l'égard de l'établissement d'enseignement est un indicateur pertinent au sein des pays. La validité interculturelle fait également l'objet d'une plus grande attention (voir la section « Validité interculturelle »).

Deux échelles de processus, le *climat de discipline en classe* et le *soutien des enseignants en classe*, sont respectivement en corrélation avec la performance des élèves et l'intérêt des élèves (voir Klieme et Rakoczy, 2003 ; Vieluf *et al.*, 2009a). Avec le temps d'apprentissage, ces échelles doivent être gardées lors des prochains cycles, à titre de variables générales de processus. Il convient toutefois de ne pas changer leur formulation lors de chaque cycle et de ne pas les associer au domaine majeur d'évaluation pour permettre l'application de modèles d'évolution.

Selon le domaine majeur d'évaluation des épreuves cognitives, des variables supplémentaires (des variables tendanciellees spécifiques au domaine) seront prises en considération. Enfin, chaque cycle a ses propres priorités, que ce soit sous la forme d'options internationales ou d'extensions aux questionnaires obligatoires.

Exploration des indicateurs du cycle PISA 2003

Les aspects suivants de l'environnement d'apprentissage en mathématiques ont été analysés sur la base des résultats du cycle PISA 2003 :

- La taille de la classe (questionnaire « Élève »).
- Le climat de discipline en cours de mathématiques (questionnaire « Élève »).
- Le soutien de l'enseignant en cours de mathématiques (questionnaire « Élève »).
- Le temps consacré à l'apprentissage des mathématiques à l'école et à domicile (questionnaire « Élève »).
- La fréquence du regroupement par aptitudes en mathématiques (questionnaire « Établissement »).
- Les activités scolaires qui promeuvent l'engagement des élèves à l'égard des mathématiques, la participation à des concours de mathématiques, par exemple (questionnaire « Établissement »).
- Les caractéristiques concernant les professeurs de mathématiques, à savoir leur niveau de qualification, leur collégialité et les politiques concernant leur évaluation (questionnaire « Établissement »).

Plusieurs résultats en rapport avec les mathématiques ont également été évalués : les stratégies d'apprentissage, l'intérêt et le plaisir, la motivation instrumentale, l'image de soi et l'anxiété.

Pour fournir des informations tendanciellees sur des résultats non cognitifs et des variables de processus et de contexte en rapport avec les mathématiques, le cycle PISA 2012 a gardé toutes les variables du cycle PISA 2003 qu'il était possible de retenir et a exclu les variables dont il est établi qu'elles présentent un biais interculturel ou n'expliquent pas de différences de résultats. Le tableau 6.2 donne un aperçu de tous les items des questionnaires « Élève » et « Établissement » du cycle PISA 2003. Dans la première rangée figurent les intrants individuels – variables démographiques et familiales – qui sont restés stables au fil des cycles⁷. Ces variables d'intrant servent à étudier la répartition des ressources et des résultats de l'éducation au sein des pays, et sont utilisées à titre de variables de contrôle dans les modèles analytiques.

Dans la deuxième rangée figurent les variables d'intrants de niveau Établissement, notamment leur statut privé ou public, leurs ressources financières, humaines et autres, et la taille et la composition de leur effectif d'élèves. Ces variables concernent des intrants à l'échelle des établissements d'enseignement, mais les décisions budgétaires, la modification des programmes de cours et les politiques de privatisation sont, en partie du moins, contrôlées au niveau du pays ou du système.

Les variables d'intrant de niveau Établissement à retenir ont été choisies sur la base des caractéristiques techniques décrites dans le rapport technique sur le cycle PISA 2003 (OCDE, 2005b) et compte tenu des analyses multiniveau du pouvoir prédictif (Vieluf *et al.*, 2009b), de la validité interculturelle (Vieluf *et al.*, 2009b) et d'une analyse de l'usage des informations obtenues *via* le questionnaire « Établissement » lors des précédents cycles PISA (Hersbach et Lietz, 2010).

Au niveau Élève, les résultats des analyses du pouvoir prédictif ont montré que la corrélation de la performance en mathématiques était relativement faible avec les préférences en matière d'apprentissage et extrêmement faible avec l'apprentissage collaboratif. Des analyses similaires réalisées sur la base des données du cycle PISA 2000 (Jehanghir et Glas, 2007) ont montré que la performance en compréhension de l'écrit n'était pas significativement corrélée avec l'apprentissage collaboratif et ne l'était que faiblement avec l'apprentissage compétitif. Ces deux échelles pourraient donc être supprimées. Parmi les trois échelles de stratégies, les stratégies de mémorisation spécifiques au domaine d'évaluation ne sont pas corrélées avec la performance en mathématiques. De plus, l'invariance interculturelle de cette



échelle n'a pu être établie – même au niveau métrique. En revanche, les stratégies d'élaboration sont fortement corrélées à la performance et l'invariance interculturelle de leur échelle est acceptable (voir la section « Indicateurs de résultats en rapport avec la culture mathématiques : stratégies, convictions et motivation »).

Au niveau Établissement, des effets relativement importants ont été observés au sujet du *sentiment d'appartenance*, du *temps consacré aux devoirs en mathématiques* et du *climat de discipline au cours de mathématiques*, des processus qui sont des variables prédictives probantes de la performance en mathématiques et de l'intérêt pour les mathématiques. Ces variables ont été retenues ; c'est le climat de discipline en classe dont le niveau d'invariance (scalaire) est le plus élevé, ce qui permet de comparer les scores moyens entre les pays.

Concernant l'*efficacité perçue* et la *perception de soi*, l'invariance scalaire des deux échelles permet d'utiliser ces variables à titre d'indicateurs de résultats non cognitifs en rapport avec les mathématiques, de suivre des tendances et de faire des analyses comparatives entre les pays. D'autres indicateurs ont toutefois été utilisés à titre expérimental pour améliorer la comparabilité interculturelle.

Rotation des items dans les extensions du questionnaire « Élève » et les instruments informatisés

Couvrir toutes les variables générales, spécifiques au domaine d'évaluation et en rapport avec des extensions thématiques serait excessif vu le temps et la place laissés aux questionnaires contextuels dans l'enquête PISA. Ce problème potentiel est particulièrement vrai pour le questionnaire « Élève », dont l'agrégation permet de générer des données à tous les niveaux de l'enquête (du niveau Élève au niveau Système). Le questionnaire « Établissement » pourrait aussi être compromis s'il était étendu, à cause du risque accru de non-réponse. Pour permettre au cycle PISA 2012 ainsi qu'aux prochains cycles de remplir de nombreux objectifs et éviter l'accroissement des données manquantes, de nouvelles méthodes s'imposent pour minimiser la charge à l'égard des répondants. Deux innovations techniques ont été introduites lors du cycle PISA 2012.

Rotation

Une rotation des items, similaire à celle appliquée aux épreuves cognitives, a été introduite dans le questionnaire « Élève » du cycle PISA 2012. Utilisée jusqu'ici dans les épreuves cognitives, la rotation permet de répartir les blocs d'items entre différents carnets. Un carnet spécifique est attribué à chaque élève, qui doit donc répondre à un nombre limité d'items cognitifs. Ensemble, les carnets couvrent la totalité des items dans les différents domaines d'évaluation. Les carnets sont attribués de façon aléatoire aux répondants dans chaque établissement d'enseignement, d'où le nom de rotation. Cette forme de rotation est courante dans les évaluations à grande échelle, qu'elles soient nationales ou internationales.

L'application d'une forme de rotation dans les questionnaires est relativement neuve dans les évaluations de l'éducation, alors que c'est une technique courante dans d'autres types d'enquête. Les implications de cette pratique pour *i)* les estimations de la compétence cognitive, *ii)* le rapport international et la présentation des tendances, *iii)* les analyses plus approfondies, *iv)* la documentation et la structure de la base de données internationale, et *v)* la logistique, ont déjà été étudiées (Berezner et Lietz, 2009). La rotation des items a été adoptée pour le questionnaire « Élève » du cycle PISA 2012 compte tenu de l'expérience acquise et d'analyses des données du cycle PISA 2006. On estime que cette approche permet d'augmenter la couverture du matériel d'un tiers.

La partie commune, sans rotation d'items, du questionnaire « Élève », doit comprendre à tout le moins les variables générales d'intrants de niveau Élève (voir la section « Répartition des contenus entre les cycles PISA »). Cette approche répond à des exigences de méthodologie et de compte rendu, et se base sur des travaux théoriques et sur des analyses des données du cycle PISA 2006 (Berezner et Lietz, 2009). Le tableau 6.3 dresse la liste des variables reprises dans la partie commune.



Tableau 6.2

Classification des questionnaires du cycle PISA 2003

	Questionnaire « Élève »	Questionnaire « Établissement »
Antécédents (Contexte personnel, familial et scolaire des élèves)	Âge, Sexe, Ascendance allochtone Langue maternelle Profession et niveau de formation des parents, richesse et patrimoine familial (= 4 indices), <u>Structure familiale</u> <u>Parcours scolaire</u> , Année d'études, Filière d'enseignement, <u>Raisons de l'assiduité</u> , Temps d'étude en dehors du cadre scolaire, Niveau de formation prévu	Taille de la communauté
Intrants scolaires (Contrôlés à l'échelle du système ou des établissements d'enseignement, selon la structure de gouvernance)		Effectifs scolarisés de sexe féminin et masculin, Pourcentage d'élèves apprenant une deuxième langue, Statut public ou privé , Années d'études, Horaires par filière Nombre d'enseignants , Nombre d'ordinateurs disponibles, Ressources financières, Ressources scolaires (3 échelles)
Processus généraux (Contrôlés à l'échelle des établissements d'enseignement et, parfois, du système)	Relations entre élèves et enseignants	Prise de décision et types d'autonomie, Politiques d'admission, Politiques d'évaluation, Options pour les élèves apprenant une deuxième langue Moral des enseignants, Comportements négatifs des enseignants, Participation des enseignants
Processus spécifiques aux mathématiques (Contrôlés à l'échelle du système ou des établissements d'enseignement)	Temps d'instruction, Temps consacré aux devoirs de mathématiques Taille des classes Climat de discipline, Soutien de l'enseignant	Regroupement par aptitudes en mathématiques Activités en mathématiques Professeurs de mathématiques : Qualification, politiques d'évaluation ; collégialité des professeurs de mathématiques
Résultats non cognitifs généraux	Attitudes à l'égard de l'établissement d'enseignement, Absentéisme, Sentiment d'appartenance à l'établissement d'enseignement	Comportements négatifs des élèves Moral des élèves Redoublement
Résultats non cognitifs spécifiques aux mathématiques	Intérêt et plaisir des mathématiques, Motivation instrumentale, Efficacité perçue, Perception de soi, Anxiété, Stratégies (contrôle, élaboration, <u>mémorisation</u>), <u>Préférences d'apprentissage</u> : apprentissage collaboratif vs apprentissage compétitif	

Remarques : Il avait été envisagé de supprimer les constructs soulignés et de retenir les constructs en **gras** lors du cycle PISA 2012. Aucune raison valable n'a motivé l'inclusion ou la suppression des constructs qui ne sont ni soulignés ni en gras.



Tableau 6.3
Partie commune du questionnaire « Élève »

N° de la question	Description
ST01	Année d'études
ST02	Filière nationale d'enseignement
ST03	Âge de l'élève
ST04	Sexe de l'élève
ST05	Fréquentation du niveau <CITE 0>
ST06	Âge au niveau <CITE 1>
ST07	Redoublement
ST08	Absentéisme – Arrivée en retard à l'école
ST09	Absentéisme – Jours d'absence non justifiée
ST115	Absentéisme – Fréquence de cours séchés
ST11	Structure familiale
ST12	Profession de la mère (CITP) ; composante de l'indice SESC
ST13	Niveau de formation de la mère (CITE) durant la scolarité ; composante de l'indice SESC
ST14	Niveau de formation de la mère (CITE) après la scolarité ; composante de l'indice SESC
ST15	Situation actuelle de la mère au regard de l'emploi ; composante de l'indice SESC
ST16	Profession du père (CITP) ; composante de l'indice SESC
ST17	Niveau de formation du père (CITE) durant la scolarité ; composante de l'indice SESC
ST18	Niveau de formation du père (CITE) après la scolarité ; composante de l'indice SESC
ST19	Situation actuelle du père au regard de l'emploi ; composante de l'indice SESC
ST20	Situation au regard de l'immigration
ST21	Âge à l'arrivée dans le pays de l'évaluation
ST25	Langue parlée en famille
ST26	Équipement général du domicile et items nationaux sur la richesse ; composante de l'indice SESC
ST27	Nombre d'articles d'équipement à domicile ; composante de l'indice SESC
ST28	Bibliothèque familiale

En outre, la partie commune du questionnaire comprend des *constructs* nécessaires pour expliquer des différences entre des groupes minoritaires majeurs. Sous l'angle de l'action publique, il s'agit de variables qui peuvent contribuer à expliquer des écarts entre les élèves selon qu'ils sont ou non issus de l'immigration.

Administration informatisée

L'informatisation des questionnaires offre une plus grande souplesse, une plus grande capacité d'adaptation, et rend l'administration et le codage plus efficaces. Ainsi, les questions de filtre sont faciles à appliquer pour guider les répondants dans le questionnaire, les incohérences peuvent être contrôlées en ligne et les données ne doivent pas être saisies manuellement. À long terme, l'informatisation des questionnaires pourrait donc aussi améliorer l'administration des instruments.

Lors du cycle PISA 2012, le questionnaire « Établissement » a été administré de façon informatisée. Cette technologie pourrait être adaptée au questionnaire « Élève » et à des options internationales lors des prochains cycles.



Tableau 6.4
Rotation A, B et C du questionnaire « Élève »

Rotation A		Rotation B		Rotation C	
Q. n°	Description	Q. n°	Description	Q. n°	Description
ST01-28	Partie commune (voir le tableau 6.3)	ST01-28	Partie commune (voir le tableau 6.3)	ST01-28	Partie commune (voir le tableau 6.3)
ST29	Motivation intrinsèque et instrumentale en mathématiques	ST42	Perception de soi en mathématiques (Q2, 4, 6, 7, 9) ; Anxiété en mathématiques (Q1, 3, 5, 8, 10)	ST53	Stratégies d'apprentissage (Contrôle vs Élaboration vs Mémorisation)
ST35	Normes subjectives	ST77	Soutien de l'enseignant en cours de mathématiques	ST55	Participation à des cours en dehors de l'école
ST37	Efficacité perçue en mathématiques	ST79	Comportement de l'enseignant : - Instruction dirigée par l'enseignant - Évaluation formative - Orientation des élèves	ST57	Nombre total d'heures consacré aux devoirs et leçons en dehors de l'école
ST43	Contrôle perçu de la performance en mathématiques	ST80	Activation cognitive en cours de mathématiques	ST61	Expérience des mathématiques appliquées (Q1, 4, 6, 8) ; Expérience des mathématiques fondamentales (Q5, 7, 9)
ST44	Raisons de l'échec en mathématiques	ST81	Climat de discipline	ST62	Mesure dans laquelle des concepts mathématiques sont familiers
ST46	Éthique en mathématiques	ST82	+ Vignettes d'ancrage	ST69	Nombre de minutes par <période de cours>
ST48	Intentions en mathématiques	ST83	Soutien du professeur de mathématiques	ST70	Nombre de <périodes de cours> par semaine
ST49	Comportement en mathématiques	ST84	+ Vignettes d'ancrage	ST71	Nombre total de <périodes de cours> par semaine
ST93	Persévérance	ST85	Gestion de la classe par le professeur de mathématiques	ST72	Taille de classe
ST94	Ouverture à la résolution de problèmes	ST86	Relations entre enseignants et élèves	ST73	Expérience des problèmes lexicaux
ST96	Stratégies de résolution de problèmes (SJT)	ST87	Sentiment d'appartenance à l'établissement d'enseignement	ST74	Expérience des tâches procédurales
ST101	Stratégies de résolution de problèmes (SJT)	ST88	Attitudes à l'égard de l'établissement d'enseignement : résultats de l'apprentissage	ST75	Expérience du raisonnement en mathématiques fondamentales
ST104	Stratégies de résolution de problèmes (SJT)	ST89	Attitudes à l'égard de l'établissement d'enseignement : activités d'apprentissage	ST76	Expérience du raisonnement en mathématiques appliquées
ST53	Stratégies d'apprentissage (Contrôle vs Élaboration vs Mémorisation)	ST91	Contrôle perçu de la réussite scolaire	ST42	Perception de soi en mathématiques (Q2, 4, 6, 7, 9) ; Anxiété en mathématiques (Q1, 3, 5, 8, 10)
ST55	Participation à des cours en dehors de l'école	ST29	Motivation intrinsèque et instrumentale en mathématiques	ST77	Soutien de l'enseignant en cours de mathématiques
ST57	Nombre total d'heures consacré aux devoirs et leçons en dehors de l'école	ST35	Normes subjectives	ST79	Comportement de l'enseignant : - Instruction dirigée par l'enseignant - Évaluation formative - Orientation des élèves
ST61	Expérience des mathématiques appliquées (Q1, 4, 6, 8) ; Expérience des mathématiques fondamentales (Q5, 7, 9)	ST37	Efficacité perçue en mathématiques	ST80	Activation cognitive en cours de mathématiques
ST62	Mesure dans laquelle des concepts mathématiques sont familiers	ST43	Contrôle perçu de la performance en mathématiques	ST81	Climat de discipline
ST69	Nombre de minutes par <période de cours>	ST44	Raisons de l'échec en mathématiques	ST82	+ Vignettes d'ancrage
ST70	Nombre de <périodes de cours> par semaine	ST46	Éthique en mathématiques	ST83	Soutien du professeur de mathématiques
ST71	Nombre total de <périodes de cours> par semaine	ST48	Intentions en mathématiques	ST84	+ Vignettes d'ancrage
ST72	Taille des classes	ST49	Comportement en mathématiques	ST85	Gestion de la classe par le professeur de mathématiques
ST73	Expérience des problèmes lexicaux	ST93	Persévérance	ST86	Relations entre enseignants et élèves
ST74	Expérience des tâches procédurales	ST94	Ouverture à la résolution de problèmes	ST87	Sentiment d'appartenance à l'établissement d'enseignement
ST75	Expérience du raisonnement en mathématiques fondamentales	ST96	Stratégies de résolution de problèmes (SJT)	ST88	Attitudes à l'égard de l'établissement d'enseignement : résultats de l'apprentissage
ST76	Expérience du raisonnement en mathématiques appliquées	ST101	Stratégies de résolution de problèmes (SJT)	ST89	Attitudes à l'égard de l'établissement d'enseignement : activités d'apprentissage
		ST104	Stratégies de résolution de problèmes (SJT)	ST91	Contrôle perçu de la réussite scolaire



GARANTIR LA VALIDITÉ DES INSTRUMENTS PISA, LEUR POUVOIR EXPLICATIF ET LEUR PERTINENCE POUR L'ACTION PUBLIQUE

Cette section présente plusieurs recommandations concernant la structure et le plan d'analyse de l'enquête PISA pour améliorer la validité et la comparabilité interculturelles, élaborer des indicateurs probants à tous les niveaux et accroître le pouvoir explicatif de l'enquête. Le perfectionnement de la structure de l'enquête sous l'angle de la méthodologie de la recherche permettra de dégager sur la base des données PISA des séries plus nombreuses et plus sophistiquées d'éclairages pertinents pour l'action publique.

La structure et les méthodes normalisées de compte rendu utilisées actuellement permettent, par exemple, de quantifier les cours de soutien proposés par les établissements d'enseignement (déclarés par les chefs d'établissement) et les cours suivis par les élèves en dehors du cadre scolaire (déclarés par les élèves et/ou leurs parents). Les responsables politiques peuvent toutefois se poser plusieurs questions à ce sujet. L'offre de cours de soutien varie-t-elle selon les établissements d'enseignement ? Les élèves ont-ils autant accès à des cours de soutien selon qu'ils sont issus d'un milieu social favorisé ou défavorisé, qu'ils sont ou non issus de l'immigration ou qu'ils suivent telle ou telle filière d'enseignement (d'après le cadre d'échantillonnage PISA) ? Existe-t-il une discrimination positive en faveur des élèves de condition modeste ? Les cours de soutien privés compensent-ils le manque de cours de soutien en milieu scolaire ? En quoi ces relations évoluent-elles au fil du temps (entre les cycles PISA) ? Où se situe mon pays par rapport à d'autres pays dans ces classements ?

Une modélisation plus sophistiquée s'impose toutefois pour répondre aux questions suivantes, et les résultats obtenus doivent être interprétés avec prudence. Le travail rémunéré empêche-t-il les élèves de participer à des cours de soutien et/ou à du tutorat ? La performance des élèves augmente-t-elle grâce à ces cours ? L'offre de cours de soutien entraîne-t-elle une augmentation de la performance moyenne à l'échelle des établissements d'enseignement, réduit-elle la variation de la performance entre les élèves inscrits dans le même établissement ou diminue-t-elle le gradient social (c'est-à-dire l'intensité de la relation entre le milieu socio-économique et la performance) au sein même des établissements d'enseignement ? Il s'agit d'hypothèses à propos de l'impact d'un traitement (le travail rémunéré, le tutorat et les cours de soutien) sur une variable (la participation à ces activités et la performance des élèves). Il faut passer par des inférences causales pour tester ces hypothèses. La validité de ces inférences dépend des variables de contrôle qui peuvent expliquer la sélection, personnelle ou non, de ces traitements (voir la section « Comprendre les modèles d'efficience dans les systèmes d'éducation »). Des analyses devraient être menées pays par pays.

La participation à des cours de soutien (imposée aux élèves ou choisie par eux) dépend vraisemblablement des pressions ou des recommandations des enseignants et des parents, et est influencée par les notes des élèves, leur sexe, leur milieu socio-économique et leur filière d'enseignement. Pour contrôler ces conditions et couvrir les variables de traitement, tous les élèves devraient être interrogés : *i*) sur la nature et la durée des cours de soutien qu'ils ont suivis en mathématiques durant l'année scolaire en cours ; *ii*) sur leurs notes à la fin de l'année scolaire précédente ; et *iii*) sur la question de savoir si leurs enseignants et/ou leurs parents leur ont recommandé, à la fin de l'année scolaire précédente, de suivre des cours de soutien. L'inclusion de ces questions dans le questionnaire « Élève » est indispensable pour faire des inférences causales à cet égard. Des effets de traitements similaires pourraient être étudiés lors de prochains cycles.

À l'échelle des établissements d'enseignement, la politique en matière d'absentéisme et son impact sur l'absentéisme des élèves, leur sentiment d'appartenance à l'établissement d'enseignement et leur performance devraient faire l'objet d'analyses. Il pourrait, par exemple, être demandé aux chefs d'établissement d'indiquer si les autorités en charge de l'éducation considéraient que l'absentéisme était un problème grave trois ans auparavant et de préciser le type de mesures mises en œuvre depuis lors.

Comment modéliser les effets des variables contextuelles sur l'apprentissage et la performance des élèves

Les précédents cycles PISA et des études sur l'efficacité de l'éducation ont analysé les effets des processus de niveau Établissement et Élève sur la performance et la motivation des élèves (par exemple, leur intérêt pour les mathématiques), après contrôle d'intrants tels que le sexe, le milieu socio-économique et l'ascendance allochtone (à l'échelle individuelle) et des mêmes intrants agrégés (à l'échelle des établissements d'enseignement). L'inclusion de ces variables de contrôle permet d'éviter de tirer des conclusions erronées de type « faux positif » à propos de relations entre deux variables qui s'expliquent en fait par une troisième variable. Dans de nombreux pays, la variation de la performance entre les élèves selon qu'ils sont ou non issus de l'immigration s'explique, par exemple, en grande partie par le fait que les élèves issus de l'immigration vivent souvent dans un milieu socio-économique défavorisé et ont, dès lors, moins de capital social et culturel. Ignorer la différence de milieu socio-économique reviendrait à



surestimer l'impact de l'ascendance allochtone. De plus, l'enquête PISA utilise des techniques de modélisation multiniveau telles que la modélisation linéaire hiérarchique pour tenir compte de la structure hiérarchique des données relatives aux élèves selon leur établissement d'enseignement et aux établissements d'enseignement selon les pays. Il est dès lors possible d'éviter de sous-estimer des erreurs types à cause du non-respect de l'hypothèse des observations indépendantes dans un échantillon, car les élèves d'un même établissement d'enseignement sont susceptibles d'avoir plusieurs caractéristiques en commun (voir, par exemple, Raudenbush et Bryk, 2002 ; Snijders et Bosker, 1999). Les modèles multiniveau permettent de surcroît l'analyse des effets à différents niveaux, c'est-à-dire les effets des variables d'intrant et de processus à l'échelle des établissements d'enseignement et des élèves.

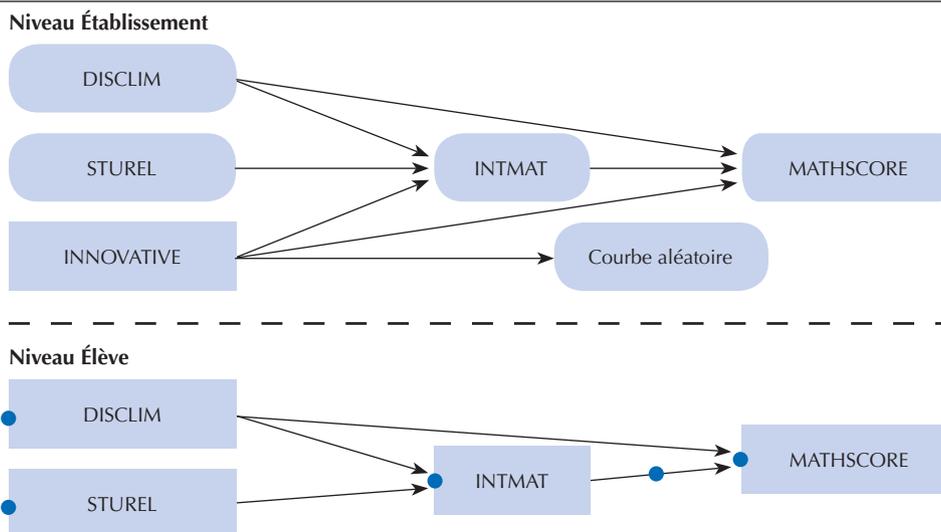
Toutefois, les études sur l'efficacité de l'éducation montrent que – après modélisation des relations entre les intrants, les processus et les extrants – trois observations supplémentaires issues de recherches antérieures devraient être prises en considération, à savoir que les effets sont souvent : *i*) non linéaires ; *ii*) atténués par d'autres variables ; et *iii*) qu'ils peuvent être indirects ou influencés par d'autres variables.

- Les effets non linéaires sont liés au fait qu'« augmentation » ne signifie pas nécessairement « amélioration » et qu'un moyen terme est parfois optimal. Monk suggère par exemple dans une étude (1994) que la relation entre la mesure dans laquelle l'enseignant maîtrise sa matière et la performance de ses élèves peut être curvilinéaire. D'une part, les enseignants sont de toute évidence incapables d'aider leurs élèves à comprendre des concepts qu'ils ne maîtrisent pas eux-mêmes et, d'autre part, les enseignants qui maîtrisent extrêmement bien leur discipline peuvent avoir des difficultés à transmettre leur savoir d'une façon que les élèves comprennent bien.
- Dans le droit fil de la théorie de l'efficacité différentielle (voir, par exemple, Kyriakides et Tsangaridou, 2004), il est important d'admettre que des relations entre variables ne sont pas nécessairement similaires dans des sous-groupes différents. Par exemple, certains éléments montrent que des élèves issus de milieux sociaux différents peuvent bénéficier de pratiques pédagogiques différentes (voir, par exemple, Brophy, 1992 ; Walberg, 1986). En pédagogie, les théories actuelles, qui s'inspirent de la conception constructiviste de l'apprentissage, se basent sur l'hypothèse que l'apprentissage des élèves dépend dans une grande mesure de processus autorégulés, qui sont modérés par des facteurs propres à l'établissement d'enseignement, à la classe et à l'enseignant. La modélisation de ces différences passe par l'analyse de ces effets d'interaction et de modération.
- Enfin, il est raisonnable de postuler que les effets sur les résultats des élèves ne sont pas tous directs. Les politiques à l'échelle des établissements d'enseignement et des systèmes ont souvent des effets plus faibles sur les résultats des élèves que des variables contextuelles de niveau Élève et des processus à l'œuvre en classe (voir, par exemple, Wang *et al.*, 1993). Ce constat peut en partie s'expliquer par le fait que les premières variables n'ont pas d'effets directs sur les élèves, mais qu'elles sont liées aux processus spécifiques à l'établissement d'enseignement et à la classe, qui ont à leur tour un effet sur la performance des élèves. De plus, des variables de niveau Établissement, telles que le climat, les normes et les valeurs ou les procédures prévues en cas de problèmes de comportement, peuvent avoir un effet direct sur des résultats non cognitifs (par exemple, la motivation pour l'apprentissage et les aspirations académiques) et le comportement des élèves (par exemple, l'absentéisme, la violence), alors que les effets de niveau Établissement sur la performance des élèves et d'autres résultats spécifiques aux matières (par exemple, l'intérêt et l'efficacité perçue) seront très vraisemblablement influencés par l'enseignement et l'apprentissage en classe.

Dans certains cas, les aspects *i*), *ii*) et *iii*) s'appliquent simultanément, par exemple lorsque les effets du climat en classe et des relations entre enseignants et élèves subissent l'influence de l'intérêt des élèves. Toutefois, l'effet de l'intérêt des élèves pour les mathématiques sur leur performance est plus important dans les établissements d'enseignement où les enseignants utilisent des pratiques pédagogiques novatrices d'activation cognitive que dans ceux où ils appliquent des pratiques plus traditionnelles. (Ce modèle est représenté dans la figure 6.2. Il s'inspire en partie des résultats d'une étude de l'enseignement des mathématiques dans deux pays par Klieme *et al.*, 2009 et Lipowsky *et al.*, 2009.) Un effet non linéaire peut être inclus s'il peut être établi que ni un climat trop laxiste, ni un climat trop strict n'est optimal pour améliorer l'intérêt des élèves ou leur performance. (Pour un exemple plus complexe basé sur l'analyse des données du cycle PISA 2003 dans un pays, voir Kaplan, 2009c et Kaplan *et al.*, 2009.)



■ Figure 6.2 ■
Exemple de modération



Remarques : DISCLIM : climat de discipline en cours de mathématiques ; STUREL : relations entre enseignants et élèves ; INNOVATIVE : stratégies pédagogiques novatrices ; INTMAT : intérêt et plaisir des mathématiques ; MATHSCORE : performance en mathématiques. Chaque point représente un paramètre estimé dans le modèle.

Les effets modérateurs, médiateurs et non linéaires peuvent être modélisés à l'aide de techniques de modélisation linéaire hiérarchique. Toutefois, la modélisation par équation structurelle se prête mieux à la modélisation des processus médiateurs (voir, par exemple, Kaplan, 2009c). Par comparaison avec la régression multiniveau, le principal avantage de la modélisation par équation structurelle réside dans le fait que les variables utilisées sont latentes et non observables, ce qui permet d'éviter les biais liés à l'erreur de mesure au niveau Élève (voir, par exemple, Muthén, 2002). Autre mérite, les mesures agrégées à l'échelle des établissements d'enseignement permettent de modéliser des effets contextuels (Lüdtke *et al.*, 2008). L'utilisation de modèles par équation structurelle améliore donc la validité des conclusions à propos de relations entre différents aspects du système d'éducation. Les modèles multiniveau par équation structurelle offrent aussi la possibilité de combiner de manière plus souple les effets modérateurs, médiateurs et non linéaires à différents niveaux, et d'analyser les variables catégorielles latentes et observables. Ainsi, une modélisation par équation structurelle pourrait être utilisée pour déterminer si l'effet du climat de discipline en classe sur la performance des élèves est similaire dans tous les pays ou s'il est plus important dans les pays où les enseignants ont moins d'autorité. Les modèles basés sur des variables latentes peuvent être appliqués pour vérifier l'hypothèse selon laquelle c'est la combinaison de stratégies d'apprentissage différentes plutôt que l'adoption fréquente de telle ou telle stratégie d'apprentissage qui a un impact sur la performance des élèves. Des modèles mixtes permettent l'identification, dans les pays, de groupes d'élèves où les relations entre le climat de discipline et la performance varient (voir, par exemple, Muthén, 2002). Ces modèles pourraient mieux refléter la réalité complexe des processus d'éducation, de leurs tenants et aboutissants.

En fait, les modèles multiniveau par équation structurelle permettraient aux responsables politiques et aux professionnels de l'éducation de mieux comprendre les relations entre : *i*) les possibilités d'apprentissage à l'échelle des établissements d'enseignement ; *ii*) l'exploitation individuelle de ces possibilités (qualité et quantité des activités d'apprentissage à l'échelle individuelle) ; et *iii*) les résultats cognitifs et motivationnels des élèves (voir, par exemple, les travaux de Kaplan *et al.*, 2009, basés sur l'analyse des données du cycle PISA 2003 dans un pays).

En résumé, il n'est pas approprié d'adopter une approche basée sur une fonction de production à l'égard des résultats de l'éducation. L'action publique et des facteurs propres aux systèmes et aux établissements d'enseignement ont des effets sur l'apprentissage des élèves *via* les processus d'enseignement, d'apprentissage et de collaboration. Les chercheurs doivent tenir compte de ces corrélations lorsqu'ils modélisent les effets à l'échelle des systèmes ou des établissements d'enseignement, et qu'ils en tirent des conclusions pertinentes pour l'action publique.



Validité interculturelle

L'objectif de l'enquête PISA est de comparer la performance des élèves et les facteurs qui la conditionnent dans un grand nombre de pays. Les études internationales permettent non seulement d'analyser les orientations politiques et d'autres intrants et processus systémiques et leurs relations avec la performance, mais également de déterminer si des conclusions de recherches pédagogiques sont transposables dans d'autres cultures. Les études internationales présentent toutefois des difficultés méthodologiques. Il faut non seulement vérifier régulièrement la fiabilité et la validité, mais également explorer les problèmes d'équivalence et de biais, deux aspects en rapport avec le temps consacré aux tâches. Les biais de *constructs* proviennent de différences réelles dans les systèmes d'éducation, les normes culturelles et l'interprétation de divers phénomènes, mais ils peuvent aussi être imputables à des erreurs de traduction ou à des problèmes de méthodologie. Pour remédier à ce problème potentiel, l'enquête PISA fait appel à des experts dans tous les pays participants. Certaines des différences dans la signification des *constructs* sont éliminées *via* des adaptations nationales, mais une modélisation par équation structurelle s'impose pour évaluer le niveau d'équivalence interculturelle et, donc, la nature des éléments qui peuvent être comparés entre les pays (moyennes, profils et corrélations). Lors des précédents cycles PISA, c'est une analyse factorielle confirmatoire multigroupe qui a été utilisée pour déterminer si le degré d'adéquation du modèle était acceptable et vérifier l'invariance des saturations factorielles. Les résultats de ces analyses montrent que les échelles PISA sont relativement équivalentes dans l'ensemble des cultures et systèmes d'éducation. Les mêmes analyses ont été effectuées lors du cycle PISA 2012, tant pour les échelles existantes que pour les nouvelles échelles. Ces dernières n'ont été retenues lors de la campagne définitive que si les résultats d'analyse étaient satisfaisants⁸.

Biais d'items

Les items sont biaisés lorsqu'ils ne se comportent pas de la même façon que d'autres items mesurant la même variable dans des groupes particuliers. Ces biais peuvent être détectés grâce à des méthodes de fonctionnement différentiel d'items qui s'inspirent de la théorie de réponse à l'item. Lors des précédents cycles PISA, ces analyses ont été réalisées lors de l'essai de terrain et de la campagne définitive. Un rapport sur les items douteux a été envoyé aux représentants des pays, et les items qui posaient problème ont été révisés (lors de l'essai de terrain) ou exclus. La même stratégie a été appliquée lors du cycle PISA 2012.

Biais de méthodes

Dans ce type de biais, le manque de comparabilité est imputable aux méthodes utilisées, c'est-à-dire à l'échantillonnage, à l'administration ou aux instruments eux-mêmes. L'un des biais courants provient d'une caractéristique des items, en l'occurrence l'utilisation d'échelles dites de Likert – courante dans les questionnaires contextuels PISA. Ces échelles sont particulièrement sensibles aux différences de style de réponse. Or, il est établi que les styles de réponse dépendent de la culture. Ce problème n'est pas sans conséquence pour la comparaison des valeurs moyennes de résultats non cognitifs et pour l'analyse des relations entre les facteurs contextuels et la performance (Hui et Triandis, 1985 ; van de Vijver *et al.*, 2008). Les styles de réponse peuvent être à l'origine de défauts d'équivalences qui peuvent être détectés *via* une analyse factorielle confirmatoire multiniveau ou une analyse du fonctionnement différentiel des items. Toutefois, des analyses supplémentaires peuvent s'imposer, car les styles de réponse ont souvent un impact similaire sur tous les items. Les différences de corrélation entre certaines échelles et la performance au niveau Élève entre les pays et au sein même de ceux-ci (voir, par exemple, van de Gaer et Adams, 2010) sont un problème connu dans l'enquête PISA. L'intérêt pour les mathématiques peut, par exemple, être en corrélation positive avec la performance en mathématiques au sein d'un pays, mais la corrélation entre les deux variables peut être fortement négative une fois comparée à l'échelle internationale. En d'autres termes, les élèves plus performants déclarent, logiquement, un intérêt plus élevé pour les mathématiques que d'autres élèves de leur pays. Dans le classement international, toutefois, des pays où la performance est plus élevée accusent des valeurs moyennes inférieures d'intérêt pour les mathématiques (par exemple, la Corée, la Finlande et le Japon), alors que des pays où la performance est moins élevée affichent des valeurs supérieures d'intérêt pour les mathématiques (par exemple, le Brésil, l'Indonésie et la Tunisie).

Une analyse plus détaillée des biais liés aux méthodes aidera à écarter les différences de scores moyens et de corrélations entre les pays qui sont dues à la variation des styles de réponse. Trois stratégies ont été utilisées lors du cycle PISA 2012 pour remédier à ce problème. En premier lieu, des analyses ont été réalisées pour déterminer si des ajustements statistiques permettraient de corriger le problème, du moins dans certaines échelles. L'adoption d'une approche de modélisation mixte (Rost *et al.*, 1997) permet, par exemple, de grouper les pays en fonction de leurs similitudes en matière de styles de réponse. La deuxième stratégie consiste à déterminer s'il est possible de remédier au problème de validité interculturelle *via* l'inclusion de divers nouveaux types d'items et d'autres formes de manipulations d'items, telles que celle proposée par Buckley (2009). À titre d'exemple, citons l'inclusion d'items inversés, d'items à choix



forcé et d'items basés sur des ancrages d'échelle différents, et le classement des items, qui permettent de manipuler l'usage des échelles de réponse et la position des items. De plus, on peut pratiquer l'alternance de formats d'items dans un questionnaire (par exemple, l'inclusion d'items basés sur une échelle de Likert à 3, 4 et 5 points dans un même questionnaire). Enfin, l'inclusion de vignettes et d'items de jugement contextuel, et la variation de la formulation des questions (par exemple, « par comparaison avec les autres élèves de votre classe » vs « dans votre pays », ou « par comparaison avec d'autres matières ») ont été analysées. D'autres procédures ont été proposées, notamment l'identification d'un groupe multilingue au sein d'un pays pour calibrer l'échelle de réponse. Comme certaines de ces méthodes requièrent la collecte de nouvelles données, elles n'ont pu être mises en œuvre que lors de l'essai de terrain.

Extension de l'échantillon pour étudier la variation au niveau Établissement (option internationale)

Le manque de données longitudinales à l'échelle des établissements d'enseignement est manifeste à la lumière de la causalité inversée (un phénomène que l'on appelle aussi déterminisme réciproque, effet simultané ou effet récursif, Scheerens et Bosker, 1997). La mise en œuvre de certaines interventions politiques (l'organisation de cours de soutien, par exemple) à cause du niveau peu élevé de performance relève, par exemple, de la causalité inversée. Dans ce cas, des données transversales peuvent donner lieu à des corrélations négatives entre la performance et des interventions politiques, même après contrôle de variables contextuelles de niveau Élève.

Scheerens et Bosker (1997) avancent l'idée que des études longitudinales à l'échelle des établissements d'enseignement permettraient de jeter des bases empiriques pour l'analyse d'effets de causalité inversée, mais ils n'en ont trouvé aucune. L'amélioration de la méthodologie des études sur l'efficacité de l'éducation (voir, par exemple, Creemers et Kyriakides, 2008) et la collecte de données d'évaluation des établissements d'enseignement par année d'études (y compris de données tendanciennes au fil du temps) ont profondément modifié ce domaine de recherche. Les chercheurs sont conscients du phénomène et savent donc que les enquêtes transversales ne permettent que jusqu'à un certain point d'expliquer pourquoi des établissements d'enseignement obtiennent des résultats plus positifs que d'autres.

Le moyen de remédier à ce problème serait de réaliser des analyses longitudinales à l'échelle des établissements d'enseignement : il faudrait échantillonner les mêmes établissements lors de cycles PISA différents pour étudier les changements qui y sont intervenus (évolution des qualifications des enseignants, de la composition socio-économique des effectifs d'élèves, des attitudes et de la performance des élèves, et variation des corrélations au fil du temps).

Extensions longitudinales de niveau Élève (recommandations pour les prochains cycles)

L'inclusion d'une composante longitudinale dans l'enquête PISA permettrait l'exploration d'un certain nombre de thématiques intéressantes, pertinentes pour l'action publique. Elle offrirait en premier lieu la possibilité d'estimer des indicateurs à valeur ajoutée, par exemple à propos de l'effet de différentes variables sur les gains de performance. Les études transversales permettent d'étudier les relations entre des variables, et les études longitudinales, d'étudier leur causalité. L'application de modèles de changement latent ou l'analyse de corrélations avec décalage permettraient, par exemple, de déterminer si l'efficacité perçue a un impact sur la performance, si la performance a un impact sur l'efficacité perçue, ou si chaque variable a un impact sur l'autre, auquel cas l'effet est réciproque.

Certains pays ont d'ores et déjà mis en œuvre une extension longitudinale de l'enquête PISA à l'échelle nationale : l'Allemagne, l'Australie (*Longitudinal Study of Australian Youth*), le Canada (*Enquête auprès des jeunes en transition*), le Danemark, la République tchèque, la Suisse (*Transitions from Education to Employment*) et l'Uruguay. Ils ont utilisé deux approches différentes :

- En Australie, au Canada, au Danemark, en Suisse et en Uruguay, ces études ont porté sur la transition entre l'enseignement secondaire et post-secondaire et/ou le monde du travail (voir, par exemple, Andersen, 2005 ; Bertschy *et al.*, 2008 ; Looker et Thiessen, 2008 ; Zoido et Gluszynski, 2009). Leurs résultats contribuent à la validation des épreuves PISA : ils montrent en effet que la performance en compréhension de l'écrit et en mathématiques est une variable prédictive probante du parcours scolaire (abandon, obtention d'un diplôme, études post-secondaires, études tertiaires, chômage, revenus, etc.). Ces études permettent d'examiner les effets de variables contextuelles et d'attitude sur des résultats ultérieurs. Il ressort, par exemple, de l'enquête menée au Canada que l'engagement académique des élèves, leurs attentes scolaires et leurs activités extrascolaires, ainsi que le soutien éducatif de leurs pairs et le soutien de leurs parents, sont des variables prédictives probantes de la poursuite de leur parcours scolaire (Looker et Thiessen, 2008 ; Thiessen, 2007). Selon l'étude réalisée en Australie, la perception individuelle des relations entre enseignants et élèves et du moral des enseignants (Curtis et McMillan, 2008) est en corrélation avec les résultats ultérieurs des élèves. Si d'autres pays mettaient cette option en œuvre, il serait possible de déterminer si les effets observés au Canada et en Australie s'appliquent à d'autres pays,



et si des caractéristiques systémiques (l'offre de programmes de « deuxième chance », par exemple) ont également un impact sur le parcours scolaire des élèves.

- Une autre approche consiste à étudier l'évolution de la performance des élèves en mathématiques durant leurs études secondaires et à analyser les corrélations entre les intrants et les processus de niveau Établissement et Élève, et la variation latente des compétences des élèves. Les études transversales permettent souvent d'observer des effets négatifs entre différentes variables de processus, tels que l'aide aux devoirs, les cours de soutien ou la coopération entre enseignants. Ces résultats peuvent en partie s'expliquer par le fait que ces indicateurs concernent les élèves en difficulté, à qui l'on donne plus de temps d'apprentissage. De même, dans les établissements d'enseignement en proie à de nombreux problèmes de discipline, les enseignants sont susceptibles de coopérer plus étroitement, car ils éprouvent un plus grand besoin de soutien. Les études longitudinales permettraient d'isoler l'impact de ces variables des impacts réels des programmes. Plus généralement, elles amélioreraient la validité des conclusions tirées des analyses concernant l'impact des politiques et des processus sur la performance des élèves, ce qui permettrait de constituer une base de connaissances plus probante pour éclairer les choix d'orientations politiques. Cette approche longitudinale a déjà été appliquée dans deux études allemandes, le Projet d'analyse du développement de la performance en mathématiques et *PISA International Plus* (PISA-I-Plus). Il ressort par exemple de ces études que la valeur moyenne de l'indicateur de l'engouement des élèves de 7^e année a un effet sensible sur l'engouement des enseignants en poste en 8^e année, qui a à son tour un effet sensible sur l'engouement des élèves de 8^e année, lequel subit l'influence de la perception de l'enthousiasme des enseignants (Frenzel *et al.*, 2009). Il en ressort également que la gestion de la classe a un effet significatif sur la performance des élèves de 10^e année en mathématiques, qui est nettement plus élevée dans les établissements d'enseignement où de nombreux enseignants sont actifs et/ou attachés à la discipline que dans les établissements d'enseignement où de nombreux enseignants sont passifs (Prenzel *et al.*, 2006).

Un autre exemple a récemment fait l'objet de débats aux États-Unis (Loveless, 2009) sur la question des élèves mal orientés (les effets négatifs de l'orientation d'élèves peu performants dans des filières exigeantes). Si les critiques formulées par Loveless (dans ce contexte spécifique) sont valables, il y aurait un effet négatif d'interaction entre le fait de suivre des cours spéciaux de physique et certains facteurs de risques (par exemple, l'ascendance allochtone ou un niveau peu élevé de performance en compréhension de l'écrit) sur la performance en mathématiques. Cette hypothèse ne peut être vérifiée comme il se doit *via* une étude transversale. Des extensions longitudinales contribueraient donc à apporter des réponses à des questions complexes en rapport avec l'action publique.

SYNTHÈSE

Ce cadre décrit les fondements des questionnaires contextuels administrés lors du cycle PISA 2012 dans le but de constituer une base de données durable pour éclairer les responsables de la politique de l'éducation et alimenter les études en matière pédagogique. Nous avons commencé par rappeler les objectifs généraux de l'enquête PISA et sa pertinence pour l'action publique. Nous avons ensuite passé en revue la littérature spécialisée par rapport à l'efficacité de l'éducation. Comme le domaine majeur d'évaluation du cycle PISA 2012 est la culture mathématique, des aspects spécifiques de l'enseignement et de l'apprentissage des mathématiques ont également été ciblés. L'essentiel de ce cadre conceptuel porte toutefois sur les fondements à retenir pour faire des questionnaires PISA des instruments durables. La structure choisie se prête à l'administration de questions générales récurrentes lors de chaque cycle et à l'administration de questions spécifiques au domaine majeur d'évaluation tous les quatre cycles pour décrire des tendances générales et spécifiques à chaque domaine. Cette structure permet aussi l'inclusion d'extensions thématiques et spécifiques, pour que l'enquête PISA puisse ajouter de nouveaux contenus ou thèmes intéressants son public. Ce cadre rappelle également des recommandations spécifiques concernant des contenus à inclure dans le cycle PISA 2012 sur la base des analyses des données PISA et d'autres études. Enfin, il se termine par des suggestions en matière de structure et d'analyse pour exploiter au mieux les points forts de l'enquête PISA et jeter des bases encore plus solides pour l'avenir.



Notes

1. Pour plus d'informations, voir la page www.oecd.org/pisa/pisainfocus/.
2. Toutes ces publications sont disponibles à l'adresse : www.oecd.org/pisa/pisaproducts/pisain-depthreports.htm.
3. Les résumés et les dossiers sont disponibles à l'adresse : www.pisaresconf09.org/.
4. Pour établir un lien entre le présent cadre d'évaluation PISA et les travaux conceptuels antérieurs de l'enquête PISA, les sections « Intrants et processus à l'échelle des systèmes », « Intrants et processus à l'échelle des établissements d'enseignement », « Intrants et processus à l'échelle des classes » et « Intrants et processus à l'échelle des élèves » s'inspirent largement du cadre conceptuel des questionnaires du cycle PISA 2009, qui a été rédigé par Hank Levin sur la base des travaux approfondis de Jaap Scheerens. Les auteurs les remercient pour leurs travaux importants et remercient les autres experts pour leur contribution aux précédents cycles PISA.
5. Voir : www.oecd.org/document/7/0,2340,en_21571361_31938349_36043527_1_1_1_1,00.html.
6. Voir EDU/PISA/GB (2005)21, p. 17, paragraphe 67.
7. Les variables qui n'ont été utilisées qu'en 2003 ou qui ont été abandonnées par la suite sont celles en rapport avec la structure familiale, le parcours scolaire et les raisons de l'assiduité scolaire.
8. L'ensemble de ces informations seront publiées dans l'ouvrage *PISA 2012 Technical Report* (à paraître).

Références

- Abedi, J., M. Courtney, S. Leon, J. Kao et T. Azzam** (2006), *English Language Learners and Math Achievement: A Study of Opportunity to Learn and Language Accommodation (CSE Report 702, 2006)*, Université de Californie, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, Californie.
- Ajzen, I.** (1991), « The Theory of Planned Behavior », *Organizational Behavior and Human Decision Processes*, vol. 50, pp. 179-211.
- Alexander, K.L., D.R. Entwisle et L.S. Olson** (2007), « Lasting Consequences of the Summer Learning Gap », *American Sociological Review*, vol. 72, pp. 167-180.
- Amiot, C. et R.Y. Bourhis** (2005), « Discrimination between Dominant and Subordinate Groups: The Positive-Negative Asymmetry Effect and Normative Processes », *British Journal of Social Psychology*, vol. 44, pp. 289-308.
- Andersen, D.** (2005), « Four Years after Basic Compulsory Education 19-year-olds on Post Compulsory Education Choices and Paths », http://www.akf.dk/udgivelses_en/2005/4aarefter_grundskolen.
- Angelone, D., U. Moser et E. Ramseier** (2009), « Instruction Time and Performance: Analyses of the Importance of Instruction Time for Competencies in Science and Mathematics based on Swiss PISA 2006 Data », dossier présenté lors de la conférence PISA, Kiel, Allemagne, 14-16 septembre, http://www.pisaresconf09.org/user_uploads/files/context/room3/Angelone_Moser_Ramseier.pdf.
- Baker, D.P.** (2009), « The Invisible Hand of World Education Culture », in G. Sykes, B. Schneider et D.N. Plank (éd.), *Handbook of Education Policy Research*, Routledge, New York, pp. 958-968.
- Baker, D.P., B. Goesling et G.K. LeTendre** (2002), « Socioeconomic Status, School Quality, and National Economic Development: A Cross-National Analysis of the 'Heyneman-Loxley effect' on Mathematics and Science Achievement », *Comparative Education Review*, vol. 46, pp. 291-312.
- Baker, D.P. et G.K. LeTendre** (2005), *National Differences, Global Similarities: World Culture and the Future of Schooling*, Stanford University Press, Stanford, Californie.
- Baker, M.L., J.N. Sigmon et M.E. Nugent** (2001), « Truancy Reduction: Keeping Students in School », *Juvenile Justice Bulletin*, Office of Juvenile Justice and Delinquency Prevention, Washington DC.
- Bandura, A.** (1977), *Social Learning Theory*, Prentice Hall, Englewood Cliffs, New Jersey.
- Baumert, J., C.H. Carstensen et T. Siegle** (2005), « Wirtschaftliche, Soziale und Kulturelle Lebensverhältnisse und Regionale Disparitäten des Kompetenzerwerbs », in PISA-Konsortium Deutschland (éd.), *PISA 2003: Der zweite Vergleich der Länder Deutschland – Was wissen und können Jugendliche?*, Waxmann Verlag, Münster.
- Baumert, J., W. Blum, M. Brunner, T. Dubberke, A. Jordan, U. Klusmann, S. Krauss, M. Kunter, K. Löwen, M. Neubrand et Y.-M. Tsai** (2008), *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV) - Dokumentation der Erhebungsinstrumente*, [Baumert et al. (2009), *Materialien aus der Bildungsforschung*, n° 83], Max-Planck-Institut für Bildungsforschung, Berlin.
- Baumert, J., M. Kunter, W. Blum, M. Brunner, T. Voss, A. Jordan, U. Klusmann, S. Krauss, M. Neubrand et Y.-M. Tsai** (2009), « Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress », *American Educational Research Journal*, vol. 47, pp. 133-180.
- Baye, A., C. Monseur et D. Lafontaine** (2009), « Institutional and Socioeconomic Factors Influencing the Reading Engagement », dossier présenté lors de la conférence PISA, Kiel, Allemagne, 14-16 septembre.
- Belfield, C. et H. Levin** (2002), « Education Privatization: Causes, Consequences, and Planning Implications », Institut international de la planification de l'éducation, UNESCO, Paris.
- Berezner, A. et P. Lietz** (2009), « Implication for Preferred Model of Rotation on Cognitive Proficiency Estimate Generation Reporting and Subsequent Analyses », dossier présenté à la réunion du groupe d'experts en charge des questionnaires, Offenbach, Allemagne, 19-21 octobre.
- Berry, J.W.** (1980), « Comparative Studies of Acculturative Stress », *International Migration Review*, vol. 21, pp. 491-511.
- Berry, J.W.** (1990), « Psychology of Acculturation », in R.W. Brislin (éd.), *Applied Cross-Cultural Psychology*, Sage, Newbury Park, Californie, pp. 232-253.
- Bertschy, K., E. Boni et T. Meyer** (2008), « Young People in Transition from Education to Labor Market. Results of the Swiss youth panel survey TREE », mise à jour de 2007, Bâle.
- Blase, J. et J. Blase** (1998), *Handbook of Instructional Leadership: How Really Good Principals Promote Teaching and Learning*, Corwin Press, Thousand Oaks, Californie.



- Blomhoj, M. et T.H. Jensen** (2007), « What's all the Fuss about Competencies? Experiences with Using a Competence Perspective on Mathematics Education to Develop the Teaching of Mathematical Modelling », in W. Blum, P.L. Galbraith, H-W. Henn et M. Niss (éd.), *Modelling and Applications in Mathematics Education*, 14 ICMI Study, Springer, New York, pp. 45-56.
- Blum, W. et D. Leiss** (2005), « How do Students and Teachers Deal with Mathematical Modelling Problems? The Example "Sugarloaf" », in *ICTMA 12 Proceedings*, Ellis Horwood, Chichester.
- Blum, W. et D. Leiss** (2007), « Investigating Quality Mathematics Teaching – the DISUM Project », in C. Bergsten et B. Grevholm (éd.), *Developing and Researching Quality in Mathematics Teaching and Learning. Proceedings of MADIF 5*, SMDF, Linköping, pp. 3-16.
- Boaler, J. et J. Greeno** (2000), « Identity, Agency and Knowing in Mathematical Worlds », in J. Bowler (éd.), *Multiple Perspectives on Mathematics Teaching and Learning*, Ablex Publishing, Westport, Connecticut.
- Borko, H., K.L. Kuffner, S.C. Arnold, L. Creighton, B.M. Stecher, F. Martinez, D. Barnes et M.L. Gilbert** (2007), *Using Artifacts to Describe Instruction: Lessons Learned from Studying Reform-Oriented Instruction in Middle School Mathematics and Science*, CSE Technical Report 705, National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, Californie.
- Bottani, N. et B. Favre** (éd.) (2001), « Open File: School Autonomy and Evaluation », *Prospects*, vol. 31.
- Britner, S.L. et F. Pajares** (2006), « Sources of Science Self-Efficacy Beliefs of Middle School Students », *Journal of Research in Science Teaching*, vol. 43, pp. 485-499.
- Bromme, R.** (1997), « Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers », in F.E. Weinert (éd.), « Reihe: Enzyklopädie der Psychologie », *Pädagogische Psychologie*, vol. 3, *Psychologie des Unterrichts und der Schule*, Hogrefe, Göttingen, pp. 177-212.
- Brookhart, S.M.** (2009), *Exploring Formative Assessment*, ASCD Publishers, Alexandria, Virginie.
- Brophy, J.** (1992) (éd.), *Planning and Managing Learning Tasks and Activities: Advances in Research on Teaching*, vol. 3, Greenwich, JAI Press, Connecticut.
- Brown, A.L.** (1994), « The Advancement of Learning », *Educational Researcher*, vol. 23, n° 8, pp. 4-12.
- Brunner, M., M. Kunter, S. Krauss, U. Klusmann, J. Baumert, W. Blum, M. Neubrand, T. Dubberke, A. Jordan, K. Löwen et Y.-M. Tsai** (2006), « Die professionelle Kompetenz von Mathematiklehrkräften: Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. Eine Zwischenbilanz des COACTIV-Projekts », in M. Prenzel et L. Allolio-Näcke (éd.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*, Waxmann Verlag, Münster, pp. 54-82.
- Buckley, J.** (2009), *Cross-National Response Styles in International Educational Assessments: Evidence from PISA 2006*, Steinhardt School of Culture, Education, and Human Development, Université de New York.
- Carnoy, M., R. Elmore et L. Siskin** (éd.) (2003), *The New Accountability: High Schools and High Stakes Testing*, Routledge Falmer, New York.
- Carroll, J.B.** (1963), « A Model of School Learning », *Teachers College Record*, vol. 64, pp. 723-733.
- Christenson, S.L.** (2004), « The Family-School Partnership: An Opportunity to Promote the Learning Competence of all Students », *School Psychology Review*, vol. 33, pp. 83-104.
- Christiansen, B. et G. Walther** (1986), « Task and Activity », in B. Christiansen, A.G. Howson et M. Otte (éd.), *Perspectives on Mathematics Education*, pp. 243-307.
- Chubb, J. et T. Moe** (1990), *Politics, Markets, and American Schools*, The Brookings Institution, Washington DC.
- Cleary, T.A.** (1968), « Test Bias: Prediction of Grades of Negro and White Students in Integrated Colleges », *Journal of Educational Measurement*, vol. 5, pp. 115-124.
- Coleman, J.** (1988), « Social Capital in the Creation of Human Capital », *American Journal of Sociology*, vol. 94, pp. 95-120.
- Common Core State Standards Initiative** (CCSSI) (2010), *Common Core State Standards*, <http://corestandards.org/>.
- Creemers, B.P.M. et L. Kyriakides** (2008), *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice, and Theory in Contemporary Schools*, Routledge, Londres.
- Curtis, D.D. et J. McMillan** (2008), *School Non-completers: Profiles and Initial Destinations. Longitudinal Surveys of Australian Youth*, Research Report 54, The Australian Council for Educational Research Ltd., Camberwell, Victoria.
- Desoete, A. et M. Veenman** (éd.) (2006), *Metacognition in Mathematics Education*, Nova Science, Hauppauge, New York.
- Flavell, J.H.** (1979), « Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry », *American Psychologist*, vol. 34, pp. 906-911.
- Frenzel, A.C., T. Goetz, O. Lüdtke, R. Pekrun et R.E. Sutton** (2009), « Emotional Transmission in the Classroom: Exploring the Relationship between Teacher and Student Enjoyment », *Journal of Educational Psychology*, vol. 101, pp. 705-716.

- Fullan, M.** (1992), *Successful School Improvement: The Implementation Perspective and Beyond*, Open University Press, Milton Keynes.
- Fullan, M.** et **S. Stiegelbauer** (1991), *The New Meaning of Educational Change*, 2^e édition, Teachers College Press, New York.
- Gamoran, A.** (1992), « The Variable Effects of High School Tracking », *American Sociological Review*, vol. 57, pp. 812-828.
- Garofalo, J.** et **F.K. Lester** (1985), « Metacognition, Cognitive Monitoring, and Mathematical Performance », *Journal for Research in Mathematics Education*, vol. 16, pp. 163-176.
- Gustafsson, J.E.** (2007), « Understanding Causal Influences on Educational Achievement through Analysis of Differences over Time within Countries », in T. Loveless (éd.), *Lessons Learned: What International Assessments Tell Us about Math Achievement*, The Brookings Institution, Washington DC, pp. 37-63.
- Hallinger, P.** et **R.H. Heck** (1998), « Exploring the Principal's Contribution to School Effectiveness: 1980 – 1995 », *School Effectiveness and School Improvement*, vol. 9, n° 2, pp. 157-191.
- Hannaway, J.** et **M. Carnoy** (éd.) (1993), *Decentralization and School Improvement: Can We Fulfill the Promise?*, Jossey-Bass, San Francisco, Californie.
- Hanushek, E.A.** et **S.G. Rivkin** (2007), « Teacher Quality », in E.A. Hanushek et F. Welch (éd.), *Handbook of the Economics of Education*, Elsevier, Amsterdam, pp. 1051-1078.
- Hanushek, E.A.** et **L. Woessmann** (2009), « Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation », *NBER Working Paper n° 14633*, National Bureau of Economic Research, Cambridge, Massachusetts.
- Hanushek, E.A.** et **L. Woessmann** (2010), « The Economics of International Differences in Educational Achievement », *IZA Discussion Paper n° 4925*, IZA, Bonn.
- Harris, A.** et **J.H. Chrispeels** (éd.) (2006), *Improving Schools and Educational Systems: International Perspectives*, Routledge, Londres.
- Heck, R.H.** et **P. Hallinger** (2005), « The Study of Educational Leadership and Management. Where does the Field Stand Today? », *Educational Management Administration and Leadership*, vol. 33, n° 2, pp. 229-244.
- Hersbach, E.** et **P. Lietz** (2010), « Towards a School Questionnaire in PISA 2012 », QEG(1006)3.doc, dossier présenté lors de la réunion du groupe d'experts en charge des questionnaires, Boston, Massachusetts, 21-23 juin.
- Hesse, H.G., K. Göbel** et **J. Hartig** (2008), « Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache », in DESI-Konsortium (éd.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch*, Weinheim, Beltz, pp. 208-230.
- Hiebert, J., R. Gallimore, H. Garnier, K.B. Givvin, H. Hollingsworth, J. Jacobs, A.M.-Y. Chui, D. Wearne, M. Smith, N. Kersting, A. Manaster, E. Tseng, W. Etterbeek, C. Manaster, P. Gonzales** et **J. Stigler** (2003), *Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study*, États-Unis, ministère fédéral de l'Éducation, National Center for Education Statistics, Washington DC.
- Hill, H.C., B. Rowan** et **D.L. Ball** (2005), « Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement », *American Educational Research Journal*, vol. 42, pp. 371-406.
- Hoover-Dempsey, K.** et **H.M. Sandler** (1997), « Why do Parents Become Involved in their Children's Education », *Review of Educational Research*, vol. 67, pp. 3-42.
- Hopkins, D.** (éd.) (2005), *The Practice and Theory of School Improvement: International Handbook of Educational Change*, Springer, Dordrecht.
- Hovey, J.D.** et **C.A. King** (1996), « Acculturative Stress, Depression, and Suicidal Ideation Among Immigrant and Second-Generation Latino Adolescents », *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 35, pp. 1183-1192.
- Hui, C.H.** et **H.C. Triandis** (1985), « Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies », *Journal of Cross-Cultural Psychology*, vol. 16, n° 2, pp. 131-152.
- Husén, T.** (1967), *International Study of Achievement in Mathematics*, vol. 2, Wiley, New York.
- Jehanghir, K.** et **C. Glas** (2007), « The Predictive Power of the Variables in the Student Questionnaires », dossier présenté à la réunion du groupe d'experts en charge des questionnaires, Arnhem, 8-9 juin.
- Kao, G.** et **M. Tienda** (1998), « Educational Aspirations of Minority Youth », *American Journal of Education*, vol. 106, n° 3, pp. 349-384.
- Kaplan, D.** (2009a), « Advances in Multilevel Latent Variable Models for PISA Data », dossier présenté lors de la conférence PISA, Kiel, Allemagne, 14-16 septembre, http://www.pisaresconf09.org/user_uploads/files/methodology/room1/Kaplan.pdf.
- Kaplan, D.** (2009b), « Causal Inference in Non-Experimental Educational Policy Research », in G. Sykes, B. Schneider et D.N. Plank (éd.), *Handbook on Education Policy Research*, Taylor and Francis, New York, pp. 139-153.
- Kaplan, D.** (2009c), *Structural Equation Modeling: Foundations and Extensions*, 2^e édition, Sage Publications, Newbury Park, Californie.



- Kaplan, D., J.-S. Kim et S.-Y. Kim** (2009), « Multilevel Latent Variable Modeling: Current Research and Recent Developments », in R.E. Millsap (éd.), *The SAGE Handbook of Quantitative Methods in Psychology*, pp. 592-613.
- Kearney, C.A.** (2008), « School Absenteeism and School Refusal Behavior in Youth: A Contemporary Review », *Clinical Psychology Review*, vol. 28, pp. 451-471.
- Klieme, E., C. Pauli et K. Reusser** (2009), « The Pythagoras Study: Investigating Effects of Teaching and Learning in Swiss and German Classrooms », in T. Janik et T. Seidel (éd.), *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, Waxmann Verlag, Münster, pp. 137-160.
- Klieme, E., G. Schümer et S. Knoll** (2001), « Mathematikunterricht in der Sekundarstufe I: Aufgabenkultur und Unterrichtsgestaltung », in Bundesministerium für Bildung und Forschung (BMBF) (éd.), *TIMSS-Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*, BMBF, Bonn, pp. 43-57.
- Klieme, E. et K. Rakoczy** (2003), « Unterrichtsqualität aus Schülerperspektive », in J. Baumert et al. (éd.), *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*, Leske und Budrich, Opladen, pp. 333-359.
- Kobarg, M., T. Seidel, M. Prenzel, B. McCrae et M. Walker** (2009), « Patterns of Science Teaching and Learning in an International Comparison », dossier présenté lors de la conférence PISA, Kiel, Allemagne, 14-16 septembre http://www.pisaresconf09.org/user_uploads/files/context/room3/Kobarg_etal.pdf.
- Koretz, D.** (2008), *What Educational Testing Really Tells Us*, Harvard University Press, Cambridge, Massachusetts.
- Krahn, H. et A. Taylor** (2005), « Resilient Teenagers: Explaining the High Educational Aspirations of Visible Minority Immigrant Youth in Canada », *Journal of International Migration and Integration*, vol. 6, n° 3/4, pp. 405-434.
- Krüger, M.L., B. Witziers et P.J.C. Sleegers** (2007), « The Impact of School Leadership on School Level Factors: Validation of a Causal Model », *School Effectiveness and School Improvement*, vol. 18, n° 1, pp. 1-20.
- Kunter, M.** (2005), *Multiple Ziele im Mathematikunterricht*, Waxmann, Münster.
- Kunter, M., Y.-M. Tsai, U. Klusmann, M. Brunner, S. Krauss et J. Baumert** (2008), « Students' and Mathematics Teachers' Perception of Teacher Enthusiasm and Instruction », *Learning and Instruction*, vol. 18, pp. 468-482.
- Kyriakides, L. et N. Tsangaridou** (2004), « School Effectiveness and Teacher Effectiveness in Physical Education », dossier présenté lors de la 85^e conférence de l'American Educational Research Association, Chicago, Illinois.
- Lee, J.** (2009), « Universals and Specifics of Math Self-Concept, Math Self-Efficacy, and Math Anxiety across 41 PISA 2003 Participating Countries », *Learning and Individual Differences*, vol. 19, n° 3, pp. 355-365.
- Lee, V.E. et D.T. Burkham** (2003), « Dropping out of High-School: The Role of School-Organization and Structure », *American Educational Research Journal*, vol. 40, n° 2, pp. 353-393.
- Lee, V.E. et J.B. Smith** (1997), « High School Size: Which Works Best and for Whom? », *Educational Evaluation and Policy Analysis*, vol. 19, pp. 205-227.
- Leithwood, K. et C. Riehl** (2005), « What we Know about Successful School Leadership », in W. Firestone et C. Riehl (éd.), *A New Agenda: Directions for Research on Educational Leadership*, Teachers College Press, New York, pp. 22-47.
- Levin, H.** (1978), « The Dilemma of Comprehensive Secondary School Reforms in Western Europe », *Comparative Education Review*, vol. 22, pp. 434-451.
- Libman, Z.** (2009), « Teacher Licensing Examinations – True Progress or Illusion? », *Studies in Educational Evaluation*, vol. 35, n° 1, pp. 7-15.
- Lie, S. et M. Kjaernsli** (2009), « Predictors for Students' Choice of Science Careers », dossier présenté lors de la conférence PISA, Kiel, Allemagne, 14-16 septembre, http://www.pisaresconf09.org/user_uploads/files/context/room3/Lie_Kjaernsli.pdf.
- Liebkind, K.** (1996), « Acculturation and Stress: Vietnamese Refugees in Finland », *Journal of Cross-Cultural Psychology*, vol. 27, pp. 161-180.
- Lipnevich, A.A., C. MacCann, S. Krumm, J. Burrus et R.D. Roberts** (à paraître), « Math Attitudes in Belarusian and US Middle School Students », *Journal of Educational Psychology*.
- Lipowsky, F., K. Rakoczy, C. Pauli, B. Drollinger-Vetter, E. Klieme et K. Reusser.** (2009), « Quality of Geometry Instruction and its Short-Term Impact on Students' Understanding of the Pythagorean Theorem », *Learning and Instruction*, vol. 19, pp. 527-537.
- Looker, E.D. et V. Thiessen** (2008), *Le système de la seconde chance : Résultats des trois cycles de l'Enquête auprès des jeunes en transition*, Direction de la politique sur l'apprentissage, Politique stratégique et de recherche, Québec, http://www.rhdcc.gc.ca/fra/publications_ressources/politique_sur_apprentissage/sp-836-04-08f/sp_836_04_08f.pdf.
- Loveless, T.** (2009), « How Well are American Students Learning? With Sections on International Assessments, the Misplaced Math Student, and Urban Schools », in *The 2008 Brown Center Report on American Education*, Brown Center on Education Policy, Brookings Institute, Washington DC.



- Lüdtke, O., H.W. Marsh, A. Robitzsch, U. Trautwein, T. Asparouhov et B. Muthén** (2008), « The Multilevel Latent Covariate Model: A New, more Reliable Approach to Group-Level Effects in Contextual Studies », *Psychological Methods*, vol. 13, pp. 203-229.
- Ma, X.** (1999), « Dropping Out of Advanced Mathematics: The Effects of Parental Involvement », *Teachers College Record*, vol. 101, pp. 60-81.
- Ma, X. et N. Kishor** (1997), « Assessing the Relationship between Attitude toward Mathematics and Achievement in Mathematics: A Meta-Analysis », *Journal for Research in Mathematics Education*, vol. 28, n° 1, pp. 26-47.
- Mau, W-C. et L. Heim Bikos** (2000), « Educational and Vocational Aspirations of Minority and Female Students: A Longitudinal Study », *Journal of Counseling and Development*, vol. 78, n° 2, pp. 186-194.
- McCluskey, C.P., T.S. Bynum et J.W. Patchin** (2004), « Reducing Chronic Absenteeism: An Assessment of an Early Truancy Initiative », *Crime and Delinquency*, vol. 50, n° 2, pp. 214-34.
- McComb, E.M. et C. Scott-Little** (2003), *A Review of Research on Participant Outcomes in After-School Programs: Implications for School Counselors*, ERIC Digest, ERIC Clearinghouse on Counseling and Student Services, ED482765, Greensboro, Caroline du Nord.
- Monk, D.H.** (1994), « Subject Matter Preparation of Secondary Mathematics and Science Teachers and Student Achievement », *Economics of Education Review*, vol. 13, n° 2, pp. 125-145.
- Mullis, I.V.S. et M.O. Martin** (2007), « TIMSS in Perspective: Lessons Learned from IEA's Four Decades of International Mathematics Assessments », in T. Loveless (éd.), *Lessons Learned. What International Assessments Tell Us about Math Achievement*, Brookings, Washington, pp. 9-36.
- Muthén, B.** (2002), « Beyond SEM: General Latent Variable Modeling », *Behaviormetrika*, vol. 29, pp. 81-117.
- National Council of Teachers for Mathematics (NCTM)** (1989), *Curriculum and Evaluation Standards for School Mathematics*, NCTM, Reston, Virginie.
- Niss, M.** (2003), « Mathematical Competencies and the Learning of Mathematics: The Danish KOM Project », in A. Gagatsis et S. Papastavridis (éd.), *3rd Mediterranean Conference on Mathematical Education*, 3-5 janvier 2003, The Hellenic Mathematical Society, Athènes, Grèce, pp. 115-124.
- Oakes, J.** (2005), *Keeping Track: How Schools Structure Inequality* (2^e éd.), Yale University, New Haven, Connecticut.
- Opendakker, M-C. et J. van Damme** (2000), « Effects of Schools, Teaching Staff and Classes on Achievement and Well-Being in Secondary Education: Similarities and Differences Between School Outcomes », *School Effectiveness and School Improvement*, vol. 11, n° 2, pp. 165-196.
- OCDE** (2001), *Connaissances et compétences : des atouts pour la vie — Premiers résultats de PISA 2000*, PISA, Éditions OCDE.
- OCDE** (2004), *Apprendre aujourd'hui, réussir demain — Premiers résultats de PISA 2003*, PISA, Éditions OCDE.
- OCDE** (2005a), *School Factors Related to Quality and Equity: Results from PISA 2000*, PISA, Éditions OCDE.
- OCDE** (2005b), *PISA 2003 Technical Report*, PISA, Éditions OCDE.
- OCDE** (2007a), *Regards sur l'éducation : Les indicateurs de l'OCDE*, Éditions OCDE.
- OCDE** (2007b), *PISA 2006 : Les compétences scientifiques : un atout pour l'avenir, volume I, Analyse des résultats*, PISA, Éditions OCDE.
- OCDE** (2008), *Regards sur l'éducation : Les indicateurs de l'OCDE*, Éditions OCDE.
- OCDE** (2009a), *Regards sur l'éducation : Les indicateurs de l'OCDE*, Éditions OCDE.
- OCDE** (2009b), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, Éditions OCDE.
- OCDE** (2010), *Résultats du PISA 2009, Volumes I, II, IV, V et VI*, PISA, Éditions OCDE.
- OCDE** (2011), *Regards sur l'éducation : Les indicateurs de l'OCDE*, Éditions OCDE.
- OCDE** (2012a), *Regards sur l'éducation : Les indicateurs de l'OCDE*, Éditions OCDE.
- OCDE** (2012b), *Untapped Skills: Realising the Potential of Immigrant Students*, PISA, Éditions OCDE.
- OCDE** (2012c), *Public and Private Schools: How Management and Funding Relate to their Socio-Economic Profile*, PISA, Éditions OCDE.
- OCDE** (2012d), *Lisons-leur une histoire ! Le facteur parental dans l'éducation*, PISA, Éditions OCDE.
- Pajares, F. et M.D. Miller** (1997), « Mathematics Self-Efficacy and Mathematical Problem Solving: Implications of Using Different Forms of Assessment », *The Journal of Experimental Education*, vol. 65, pp. 213-228.
- Perreault, S. et R.Y. Bourhis** (1999), « Ethnocentrism, Social Identification and Discrimination », *Personality and Social Psychology Bulletin*, vol. 25, pp. 92-103.



Prenzel, M., J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost et U. Schiefele (éd.) (2006), *PISA 2003 : Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*, Waxmann Verlag, Münster.

Purves, A.C. (1987), « The Evolution of the IEA: A Memoir », *Comparative Education Review*, vol. 31, n° 1, pp. 10-28.

Rakoczy, K., A. Buff et F. Lipowsky (2005), « Teil 1: Befragungsinstrumente », in E. Klieme et al. (éd.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie Unterrichtsqualität, Lernverhalten und mathematisches Verständnis*, GPF, Francfort.

Rakoczy, K., E. Klieme et C. Pauli (2008), « Die Bedeutung der wahrgenommenen Unterstützung motivationsrelevanter Bedürfnisse und des Alltagsbezugs im Mathematikunterricht für die selbstbestimmte Motivation », *Zeitschrift für Pädagogische Psychologie*, vol. 22, n° 1, pp. 25-35.

Raudenbush, S.W. et A.S. Bryk (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, Thousand Oaks, Californie.

Rost, J., C.H. Carstensen et M. von Davier (1997), « Applying the Mixed Rasch Model to Personality Questionnaires », in J. Rost et R. Langeheine (éd.), *Applications of Latent Trait and Latent Class Models in the Social Sciences*, Waxmann, New York, pp. 324-332.

Rowan, B., E. Camburn et R. Correnti (2004), « Using Teacher Logs to Measure the Enacted Curriculum in Large-Scale Surveys: Insights from the Study of Instructional Improvement », *Elementary School Journal*, vol. 105, pp. 75-102.

Ryan, R.M. et E.L. Deci (2000), « Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions », *Contemporary Educational Psychology*, vol. 25, pp. 54-67.

Rychen, D.S. et L.H. Salganik (éd.) (2003), *Key Competencies for a Successful Life and a Well-Functioning Society*, Hogrefe and Huber Publishers, Göttingen.

Scheerens, J. (2000), « Improving School Effectiveness », série *Principes de la planification de l'éducation*, IIEP, vol. 68, UNESCO, Paris.

Scheerens, J. et R.J. Bosker (1997), *The Foundations of Educational Effectiveness*, Pergamon, Oxford.

Schmid, C.L. (2001), « Educational Achievement, Language-Minority Students, and the New Second Generation », *Sociology of Education* (hors série), pp. 71-87.

Schmidt, W.H. et C.C. McKnight (1998), « What Can We Really Learn from TIMSS? », *Science*, vol. 282, pp. 1831-1839.

Schmidt, W.H., C.C. McKnight, R.T. Houang, H.C. Wang, D.E. Wiley, L.S. Cogan et R.G. Wolfe (2001), *Why Schools Matter: A Cross-National Comparison of Curriculum and Learning*, Jossey Bass, San Francisco, Californie.

Schmidt, W.H. et R.T. Houang (2007), « Lack of Focus in the Mathematics Curriculum: Symptom or Cause », in T. Loveless (éd.), *Lessons Learned: What International Assessments Tell Us about Math Achievement*, Brookings, Washington, pp. 65-84.

Schmidt, W.H. et A. Maier (2009), « Opportunity to Learn », in G. Sykes, B. Schneider et D.N. Plank (éd.), *Handbook of Education Policy Research*, Routledge, New York, pp. 541-559.

Schmidt, W.H. et C. McKnight (1995), « Surveying Educational Opportunity in Mathematics and Science: An International Perspective », *Educational Evaluation and Policy Analysis*, vol. 17, n° 3, pp. 337-353.

Schneider, W. et C. Artelt (à paraître), « Metacognition and Mathematics Education », *The International Journal on Mathematics Education*, ZDM.

Schoenfeld, A.H. (éd.) (1987), *Cognitive Science and Mathematics Education*, Erlbaum, Hillsdale, New Jersey.

Schoenfeld, A.H. (1992), « Learning to Think Mathematically: Problem Solving, Metacognition, and Sense-Making in Mathematics », in D. Grouws (éd.), *Handbook for Research on Mathematics Teaching and Learning*, MacMillan, New York, pp. 334-370.

Schoenfeld, A.H. (2006), « Mathematics Teaching and Learning », in P.A. Alexander et I.H. Winne (éd.), *Second Handbook of Educational Psychology*, Routledge, Londres.

Shulman, L. (1987), « Knowledge and Teaching: Foundations of the New Reform », *Harvard Educational Review*, vol. 57, n° 1, pp. 1-22.

Slavin, R.E. (1990), « Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis », *Review of Educational Research*, vol. 60, n° 3, pp. 471-499.

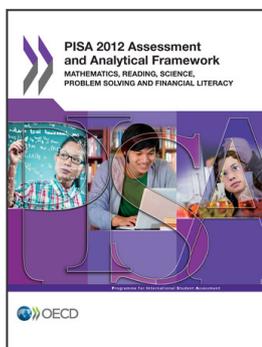
Snijders, T. et R. Bosker (1999), *Multilevel Analysis*, Sage Publications, Londres.

Stanat, P. (2006), « Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft », in J. Baumert, P. Stanat et R. Watermann (éd.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit; vertiefende Analysen im Rahmen von PISA 2000*, VS Verl. für Sozialwiss, Wiesbaden, pp. 189-219.

Stanat, P. et G. Christensen (2006), *Where Immigrant Students Succeed - A Comparative Review of Performance and Engagement in PISA 2003*, Éditions OCDE.



- Stein, M.K., R.A. Engle, M.S. Smith et E.K. Hughes** (2008), « Orchestrating Productive Mathematical Discussions: Five Practices for Helping Teachers Move Beyond Show and Tell », *Mathematical Thinking and Learning*, vol. 10, pp. 315.
- Stevens, F.** (1993), « Applying an Opportunity-to-Learn Conceptual Framework to the Investigation of the Effects of Teaching Practices via Secondary Analyses of Multiple-Case-Study Summary Data », *Journal of Negro Education*, vol. 62, n° 3, pp. 232-248.
- Stigler, J.W. et J. Hiebert** (1999), *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*, Free Press, New York.
- Sui-Chu, H. et J.D. Willms** (1996), « Effects of Parental Involvement on Eighth-Grade Achievement », *Sociology of Education*, vol. 69, pp. 126-141.
- Teddlie, C. et D. Reynolds** (éd.) (2000), *The International Handbook of School Effectiveness Research*, Routledge, New York.
- Thiessen, V.** (2007), *The Impact of Factors on Trajectories that Lead to a High School Diploma and to Participation in Post Secondary Education*, Ressources humaines et Développement des compétences, Canada, Québec.
- Turner, S.L., J.C. Steward et R.T. Lapan** (2004), « Family Factors Associated with Sixth Grade Adolescents' Math and Science Career Interests », *Career Development Quarterly*, vol. 53, n° 1, pp. 41-52.
- Van de Gaer, E. et R. Adams** (2010), « The Modeling of Response Style Bias: An Answer to the Attitude-Achievement Paradox? », dossier présenté à la conférence annuelle de l'American Educational Research Association, Denver, Colorado, 30 avril-4 mai.
- Van de Vijver, F.J.R., D.A. Van Hemert et Y.H. Poortinga** (éd.) (2008), *Multilevel Analysis of Individuals and Cultures*, Erlbaum, New York.
- Van Landeghem, G., B. De Fraine et J. Van Damme** (2005), « The Consequence of Ignoring a Level of Nesting in Multilevel Analysis: A Comment », *Multivariate Behavioral Research*, vol. 40, pp. 423-434.
- Vieluf, S., J. Lee et P. Kyllonen** (2009a), « The Predictive Power of Variables from the PISA 2003 Student Questionnaire », dossier présenté lors de la réunion du groupe d'experts en charge des questionnaires, Offenbach, Allemagne, 19-21 octobre.
- Vieluf, S., J. Lee et P. Kyllonen** (2009b), « The Cross-Cultural Validity of Variables from the PISA 2003 Student Questionnaire », dossier présenté lors de la réunion du groupe d'experts en charge des questionnaires, Offenbach, Allemagne, 19-21 octobre.
- Walberg, H.J.** (1986), « *Syntheses of Research on Teaching* », in M.C. Wittrock (éd.), *Handbook of Research on Teaching* (3^e édition), Macmillan, New York, pp. 214-229.
- Wang, M.C., G.D. Haertel et H.D. Walberg** (1993), « Toward a Knowledge Base for School Learning », *Review of Educational Research*, vol. 63, n° 3, pp. 249-294.
- Wayne, A. et P. Youngs** (2003), « Teacher Characteristics and Student Achievement Gain: A Review », *Review of Educational Research*, vol. 73, pp. 89-122.
- Wigfield, A., J.S. Eccles et D. Rodriguez** (1998), « The Development of Children's Motivation in School Contexts », in A. Iran-Nejad et P.D. Pearson (éd.), *Review of Research in Education*, American Educational Research Association, vol. 23, Washington DC.
- Wilmers, N., D. Enzmann, D. Schaefer, K. Herbers, W. Greve et P. Wetzels** (2002), *Jugendliche in Deutschland zur Jahrtausendwende: Gefährlich oder gefährdet? Ergebnisse wiederholter, repräsentativer Dunkelfelduntersuchungen zu Gewalt und Kriminalität im Leben junger Menschen 1998 – 2000*, Nomos Verlagsgesellschaft, Baden-Baden, Allemagne.
- Witziers, B., R.J. Bosker et M.L. Krüger** (2003), « Educational Leadership and Student Achievement: The Elusive Search for an Association », *Educational Administrative Quarterly*, vol. 39, n° 3, pp. 398-425.
- Woessmann, L.** (2006), « The Complementarity of Central Exams and School Autonomy: Economic Theory and International Evidence », in E. Gori, D. Vidoni, E. Hanushek et C. Glenn (éd.), *Institutional Models in Education: Legal Framework and Methodological Aspects for a New Approach to the Problem of School Governance*, Wolf Legal Publishers, Nimègue, pp. 47-65.
- Yasushi, O.** (2009), « Comparison of Attitudes toward Science between Grade 9 and 10 Japanese Students by Using the PISA Questions and Its Implications on Science Teaching in Japan », dossier présenté lors de la réunion du groupe d'experts en charge des questionnaires, Offenbach, Allemagne, 19-21 octobre.
- Zoido, P. et T. Gluszynski** (2009), « Pathways to Success: How Knowledge and Skills at Age 15 Shape Future Life Outcome », dossier présenté à la conférence PISA, Kiel, Allemagne, 14-16 septembre.
- Zuzovsky, R.** (2009), « Teachers' Qualifications and Their Impact on Student Achievement: Findings from TIMSS 2003 Data for Israel », in M. von Davier et D. Hastedt (éd.), *IERI Monograph Series Issues and Methodologies in Large-Scale Assessments*, IER Institute, Hambourg/Princeton, Allemagne/États-Unis, vol. 2, pp. 37-62.



Extrait de :

PISA 2012 Assessment and Analytical Framework Mathematics, Reading, Science, Problem Solving and Financial Literacy

Accéder à cette publication :

<https://doi.org/10.1787/9789264190511-en>

Merci de citer ce chapitre comme suit :

OCDE (2013), « Cadre conceptuel des questionnaires contextuels du cycle PISA 2012 », dans *PISA 2012 Assessment and Analytical Framework : Mathematics, Reading, Science, Problem Solving and Financial Literacy*, Éditions OCDE, Paris.

DOI: <https://doi.org/10.1787/9789264190559-8-fr>

Cet ouvrage est publié sous la responsabilité du Secrétaire général de l'OCDE. Les opinions et les arguments exprimés ici ne reflètent pas nécessairement les vues officielles des pays membres de l'OCDE.

Ce document et toute carte qu'il peut comprendre sont sans préjudice du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région.

Vous êtes autorisés à copier, télécharger ou imprimer du contenu OCDE pour votre utilisation personnelle. Vous pouvez inclure des extraits des publications, des bases de données et produits multimédia de l'OCDE dans vos documents, présentations, blogs, sites Internet et matériel d'enseignement, sous réserve de faire mention de la source OCDE et du copyright. Les demandes pour usage public ou commercial ou de traduction devront être adressées à rights@oecd.org. Les demandes d'autorisation de photocopier une partie de ce contenu à des fins publiques ou commerciales peuvent être obtenues auprès du Copyright Clearance Center (CCC) info@copyright.com ou du Centre français d'exploitation du droit de copie (CFC) contact@cfcopies.com.