



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



Bulletin d'information
n° 54
Septembre 2018
UIS/2018/ED/IP/54

Les écarts en matière d'apprentissage : Utiliser les données pour orienter la politique éducative



UNESCO

L'Acte constitutif de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) a été adopté par 20 pays lors de la Conférence de Londres en novembre 1945 et est entré en vigueur le 4 novembre 1946.

L'UNESCO compte actuellement 195 États membres et 11 membres associés.

L'objectif premier de l'UNESCO est de contribuer au maintien de la paix et de la sécurité dans le monde en resserrant, par l'éducation, la science et la culture, la collaboration entre nations afin d'assurer le respect universel de la justice, de la loi, des droits de l'homme et des libertés fondamentales pour tous, sans distinction de race, de sexe, de langue ou de religion, que la Charte des Nations Unies reconnaît à tous les peuples.

L'UNESCO a cinq fonctions principales, inscrites dans son mandat : 1) des études prospectives sur l'éducation, la science, la culture et la communication dans le monde de demain; 2) le progrès, le transfert et le partage des connaissances par des activités de recherche, de formation et d'enseignement; 3) des actions normatives en vue de la préparation et de l'adoption d'instruments internationaux et de recommandations réglementaires; 4) l'expertise par le biais de la coopération technique avec les États membres, en faveur de leurs projets et politiques de développement; 5) l'échange d'informations spécialisées.

Institut de statistique de l'UNESCO

L'Institut de statistique de l'UNESCO (ISU), qui forme le bureau de statistique de l'UNESCO, agit comme organisation dépositaire de l'ONU de statistiques à l'échelle mondiale dans les domaines de l'éducation, de la science et de la technologie, de la culture et des communications.

L'ISU a été mis sur pied en 1999. Il a été créé pour améliorer le programme statistique de l'UNESCO et pour élaborer et diffuser en temps opportun des statistiques fiables, précises et pertinentes en matière politique qu'exigent les environnements social, politique et économique d'aujourd'hui, en constante évolution et de plus en plus complexes.

Le présent rapport a été écrit par J. Douglas Willms, Université du Nouveau-Brunswick et président de The Learning Bar Inc., doug.w@thelearningbar.com

Publié en 2018 par :

Institut de statistique de l'UNESCO

C.P. 6128, Succursale Centre-Ville

Montréal (Québec) H3C 3J7

Canada

Tél : +1 514-343-6880

Courriel : uis.publications@unesco.org

<http://www.uis.unesco.org>

Réf : UIS/2018/ED/IP54

© UNESCO-ISU 2018

Cette publication est disponible en accès libre sous la licence Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). En utilisant le contenu de cette publication, les utilisateurs acceptent de se conformer aux conditions d'utilisation du service d'archive des publications en accès libre de l'UNESCO (<https://fr.unesco.org/open-access/terms-use-ccbysa-fr>).

Les appellations employées dans cette publication et la présentation des données qui y figurent n'impliquent de la part de l'UNESCO aucune prise de position quant au statut juridique de tout pays, territoire, ville ou zone ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.



Les idées et opinions exprimées dans cette publication sont celles des auteurs. Elles ne représentent pas nécessairement les vues de l'UNESCO et n'engagent à rien l'Organisation.



Remerciements

J'aimerais remercier Danielle Durepos, Elizabeth Fairbairn, Alma Lopez et Lucia Tramonte pour leurs commentaires sur les ébauches de ce rapport, Jennifer Morgen, qui a fait la programmation des figures utilisées dans celui-ci, et les autres membres du personnel de The Learning Bar Inc. qui ont soutenu ce travail à diverses étapes. Bon nombre des idées avancées dans le présent rapport sont le fruit d'une longue collaboration avec Lucia Tramonte, qui a aidé à clarifier un certain nombre de concepts afin qu'ils puissent être appliqués plus directement à la recherche menée pour le Programme de développement durable de l'ONU.

Le modèle de prospérité éducative a été élaboré avec la collaboration de leaders de l'éducation au Canada et en Australie, et a ensuite été modifié pour le PISA pour le développement de l'OCDE. Pendant le stade de développement, j'ai reçu de nombreux commentaires judicieux de la part d'éducateurs et de directeurs de plusieurs territoires. Je suis reconnaissant des contributions de Michael Ward et du personnel de l'OCDE, des membres du Groupe d'experts sur les questionnaires et du Groupe consultatif international, et des analystes et des gestionnaires de projet nationaux des huit pays qui ont participé au PISA pour le développement.

J'aimerais remercier Eduardo Backhoff, David Berliner, Harvey Sanchez, Valérie Tehio et Servaas Van der Berg, qui ont révisé attentivement le rapport et qui ont fait des observations précieuses.

Enfin, je voudrais Silvia Montoya, directrice de l'Institut de statistique de l'UNESCO, qui a soutenu cette recherche et qui m'a encouragé à écrire ce rapport.



Table des matières

Remerciements	4
Introduction	6
I. Prospérité éducative : une approche du parcours de vie	11
S'épanouir	11
Résultats de prospérité	12
Bases de la réussite	12
Quatre façons de cumuler la réussite	13
Égalité, équité et accès	14
Le rôle du SSÉ	17
II. Deux périodes de transition critiques pour le développement des compétences en littératie	19
Entrée à l'école primaire	19
De « apprendre à lire » à « lire pour apprendre »	22
III. Politiques sur les résultats des élèves et l'égalité	27
Où en sommes-nous?.....	27
Qui est vulnérable?	31
Où sont les enfants vulnérables?.....	35
IV. Politiques sur les stratégies et leur mise en œuvre.....	42
Renforcer les bases de la réussite	47
Modifier les caractéristiques structurelles des écoles	52
V. Faire le suivi de la prospérité éducative	54
Établir des objectifs.....	54
Un exemple de programme de suivi	57
Nouvelles orientations	57
Références	60
Annexe 1. Taux de croissance annuelle de la performance en lecture, PISA 2000 à 2015.....	68
Annexe 2. Abrégé des statistiques essentielles en politique éducative	69
Figure 1. Taux de croissance annuelle de la performance en lecture, PISA 2000 à 2015	9
Figure 2. Le cadre de prospérité éducative	11
Figure 3. Quatre façons de cumuler la réussite	13
Figure 4. Égalité, équité et effets de l'école sur la prospérité éducative	15
Figure 5. Pourcentage des élèves ayant des compétences en lecture de niveau 2 ou inférieur à l'âge de 15 ans par rapport au revenu national brut (méthode Atlas en dollars américains).....	18
Figure 6. Scores pour le domaine des compétences cognitives et langagières à l'Évaluation de la petite enfance, Uruguay, 2017	21
Figure 7. Scores en lecture en Australie, NAPLAN 2017	24



Figure 8. Distribution de la performance en lecture à l'âge de 15 ans en République dominicaine, au Mexique et en Suède	28
Figure 9. Relation entre la performance moyenne en lecture et l'asymétrie	29
Figure 10. Gradient socioéconomique de la performance en lecture pour le Mexique	32
Figure 11. Gradients socioéconomiques de la performance en lecture selon la langue maternelle au Mexique.....	34
Figure 12. Profil des écoles pour la performance en lecture au Mexique	35
Figure 13. Profil des écoles pour les élèves ayant de faibles compétences en lecture (niveau 1 et inférieur) au Mexique	36
Figure 14. Concentration d'élèves avec de faibles compétences en lecture dans les écoles du Mexique	39
Figure 15. Effets d'une stratégie universelle avec une ampleur de l'effet de 0,50	44
Figure 16. Effets d'une stratégie ciblant la performance avec une ampleur de l'effet de 0,50 pour les élèves ayant des compétences en lecture de niveau 1 ou inférieur	44
Figure 17. Effets d'une stratégie ciblant les risques avec une ampleur de l'effet de 0,50 pour les élèves dont le SSÉ est de -2 ou moins	45
Figure 18. Effets d'une stratégie compensatoire qui augmente le SSÉ des élèves de faible SSÉ par 0,25 d'un écart type	46
Figure 19. Effets d'une stratégie de redistribution visant à replacer les élèves des écoles de faible SSÉ dans des écoles régulières	47
Figure 20. Cadre de prospérité éducative pour le PISA pour le développement.....	48
Figure 21. Ressources des écoles au Mexique.....	51
Figure 22. Domaines et mesures pour un système de suivi national.....	58



Introduction

Le niveau de compétence en lecture des élèves varie grandement d'un pays à l'autre, et même d'un pays riche à l'autre. Ce niveau de compétence varie également beaucoup d'une école à l'autre au sein d'un même pays. Ce sont les deux conclusions du rapport *Les écarts en matière d'apprentissage : Dix questions stratégiques sur le rendement et l'équité des écoles et des systèmes d'éducation* (Willms, 2006). On parle d'écarts en matière d'apprentissage parce que les analyses ont révélé que le lien entre le niveau de compétence en lecture des élèves et le statut socioéconomique (SSÉ) de leur famille est très différent d'un pays à l'autre. Ce lien est un excellent indicateur des inégalités, ou des écarts en matière d'apprentissage, entre les élèves de faible SSÉ et de SSÉ élevé. Le niveau de compétence en lecture des élèves est généralement plus élevé dans les pays où il y a moins d'inégalités. Ces différences entre les pays et entre les écoles d'un même pays sont évidentes à la fin de la 4^e année et s'accroissent ensuite au fil des années.

Le rapport de 2006 est fondé sur des analyses de données du Programme international de recherche en lecture scolaire (PIRLS) (2001) et du Programme international pour le suivi des acquis des élèves (PISA) (2000). Le PIRLS est une évaluation des compétences en lecture des élèves de 4^e année du primaire qui est réalisée tous les cinq ans depuis 2001 sous les auspices de l'Association internationale pour l'évaluation du rendement scolaire (IEA). Le PISA est une évaluation des connaissances et des compétences pour la vie quotidienne réalisée auprès des jeunes de 15 ans par l'Organisation de coopération et de développement économiques (OCDE). Le PISA s'intéresse aux compétences en lecture, en mathématiques et en sciences dont les élèves auront besoin au quotidien dans leurs études postsecondaires ou sur le marché du travail. Il est réalisé tous les trois ans depuis 2000.

Le rapport *Les écarts en matière d'apprentissage* utilisait des gradients socioéconomiques comme cadre pour aborder diverses questions de politique éducative. Un gradient socioéconomique – ou niveau d'apprentissage – décrit le lien entre un résultat scolaire ou social et le SSÉ des personnes d'un territoire donné (p. ex. une école, une région ou un pays) (Willms, 2003a). Le principal objectif d'une politique éducative est d'améliorer les résultats des élèves et de réduire les inégalités associées au contexte familial. En d'autres termes, d'augmenter et d'uniformiser le niveau d'apprentissage. Le rapport démontre qu'une analyse détaillée des gradients socioéconomiques qui tient compte de la structure hiérarchique du système scolaire peut fournir de l'information susceptible d'aider les décideurs à mieux intervenir pour améliorer les compétences des élèves et réduire les inégalités. Le rapport aborde également la question de la valeur ajoutée : « La variation inter-écoles observée est-elle attribuable aux ressources des écoles et aux politiques et pratiques des écoles et des groupes? » (Willms, 2006, p. 54). Les analyses ont montré que, lorsque les contextes familiaux sont pris en compte, le lien entre la performance des élèves et les mesures des pratiques des écoles et des groupes est faible et le plus souvent non statistiquement significatif.

Le présent rapport est une suite au rapport de 2006. Il est basé sur le cadre de prospérité éducative, lequel peut être utilisé pour évaluer si les familles, les communautés et les institutions publiques réussissent à développer les compétences cognitives des enfants et à assurer leur bien-être social, émotionnel, physique et spirituel. Ce cadre d'évaluation adopte une approche du parcours de vie qui tient compte des processus



qui déterminent les résultats des enfants de la conception jusqu'à la fin de l'adolescence. Celui-ci comprend un ensemble principal de résultats (les « résultats de prospérité ») pour chacun des six stades de développement et un ensemble de facteurs familiaux, institutionnels et communautaires (les « bases de la réussite ») qui influencent ces résultats. Le cadre fait également une distinction entre les notions d'égalité et d'équité, et établit des liens explicites entre le suivi des systèmes d'éducation et la politique nationale et locale. Le cadre de prospérité éducative est décrit plus en détail dans un document séparé (Willms, 2018a).

Les études internationales de grande envergure qui orientent la politique éducative ont surtout adopté deux approches. La première consiste à recueillir des données sur une myriade de facteurs en lien avec l'école et le groupe, et à déterminer les relations entre ces facteurs et les résultats scolaires. La recherche des « effets de l'école » est une caractéristique centrale des études internationales de grande envergure, comme celles du PISA et de l'IEA. Les cadres contextuels sont basés sur un paradigme de cause à effet qui tente de faire ressortir les facteurs associés à l'élève, à la famille, au groupe et à l'école qui expliquent le mieux la performance des élèves. Les analyses de données sont basées sur des modèles de régression multiniveaux qui examinent les relations entre un résultat d'un élève, par exemple en lecture, et une longue liste de facteurs associés à l'école et au groupe. Souvent, ces modèles statistiques sont estimés séparément pour chaque pays parce qu'on pense que les facteurs associés à l'école et au groupe qui sont pertinents pour les compétences des élèves ne sont pas les mêmes d'un pays à l'autre. Les résultats de ces analyses viennent ensuite soutenir diverses politiques nationales.

Dans la deuxième approche, les pays comparent leurs résultats à ceux d'autres pays. Les facteurs qui sont considérés comme pertinents pour la réussite des élèves sont regroupés par thèmes, les principaux étant les ressources des écoles, la responsabilité, la gouvernance, les méthodes d'enseignement et la scolarité sélective. Un pays peut chercher à savoir si ses politiques sont différentes de celles d'autres pays ayant des contextes socioéconomiques similaires. Toutefois, la communauté politique internationale a surtout tendance à s'intéresser aux histoires de réussite. Par exemple, les excellents résultats de la Finlande à la première enquête du PISA ont donné lieu à une multitude d'hypothèses pour expliquer cette réussite, la majorité pointant vers l'expertise des enseignants finnois et l'approche adoptée pour l'évaluation des élèves (Grek, 2009; Simola, 2005).

Les deux approches sont problématiques. Le principal problème est que la performance des élèves aux tests du PIRLS à la fin de la 4^e année ou à ceux du PISA à l'âge de 15 ans est le résultat d'un nombre incalculable de facteurs ayant une incidence sur le développement des enfants, depuis la conception jusqu'au moment de l'évaluation. Les compétences cognitives et langagières des élèves à leur entrée en 1^{re} année sont de forts prédicteurs des compétences qu'ils auront en lecture trois ans plus tard (Scarborough, 1989; Schatschneider *et al.*, 2004). Les compétences en lecture à la fin du primaire sont un fort prédicteur des compétences en lecture à l'âge de 15 ans (Adlof, Catts et Lee, 2010). Il ne faut donc pas s'attendre à ce que les mesures des pratiques des écoles et des groupes (lesquelles sont dérivées des réponses à des questionnaires remplis au même moment que les tests) aient de forts liens avec la performance en lecture.



Par ailleurs, les mesures des facteurs scolaires clés ayant une incidence sur la performance des élèves ont tendance à être intercorrélées et sont fortement corrélées au SSÉ moyen de l'école. Il est pratiquement impossible d'isoler les « effets de l'école » attribuables à des ressources ou à des processus en particulier avec une étude transversale (Raudenbush et Willms, 1995). La deuxième approche est problématique pour les mêmes raisons. Il est possible qu'un pays ait des résultats supérieurs à ceux d'un autre parce qu'il soutient le développement des enfants pendant leurs premières années de vie. Dans un tel cas, toute comparaison des effets des politiques scolaires peut produire de faux résultats. En d'autres termes, il est impossible d'établir des liens de cause à effet à partir des résultats d'études nationales ou internationales.

Il n'y a eu aucune amélioration des compétences en lecture des élèves au cours des 15 dernières années. Il s'agit de la raison la plus convaincante pour adopter une nouvelle approche pour l'utilisation des données qui orientent la politique éducative. La **figure 1** montre les taux de croissance annuelle pour les 28 pays qui ont participé au PISA entre 2000 et 2015. Les scores en lecture à l'évaluation de 2000 ont été échelonnés pour obtenir une moyenne de 500 et un écart type de 100 pour tous les pays de l'OCDE. Les échelles ont été égalisées pour chaque cycle afin que les changements puissent être estimés à partir d'une échelle commune.¹ Les points verts indiquent que le taux de croissance d'un pays est statistiquement significatif. Les résultats détaillés sont présentés sous la forme d'un tableau à l'annexe 1.

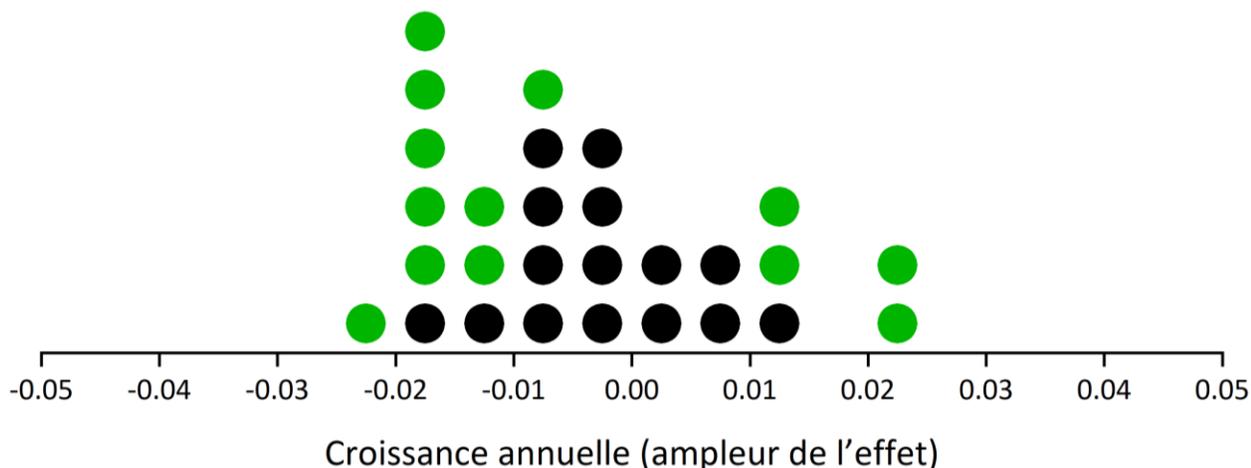
En moyenne, le taux de croissance annuelle est légèrement négatif, soit -0,004 % d'un écart type. Les taux de croissance annuelle ne sont pas statistiquement significatifs pour 15 des 28 pays. Parmi les taux de croissance annuelle qui sont significatifs, neuf sont négatifs et quatre sont positifs. Pour tous les pays, les taux de croissance annuelle sont inférieurs à 2,5 % d'un écart type (positivement ou négativement). Cela ne représente que 2,5 points sur l'échelle du PISA, laquelle a une moyenne de 500 et un écart type de 100. Pour mettre ces chiffres dans un contexte plus large, il faut savoir qu'une intervention en lecture bien conçue peut améliorer la performance des élèves de 50 ou 60 points, ce qui correspond à une ampleur de l'effet de 0,50 à 0,60. Cette estimation est basée sur la synthèse de 50 méta-analyses d'interventions en lecture de Hattie (2009). L'ampleur de l'effet est de 0,51 en moyenne. Pour les 14 études réalisées pour la méthode phonétique, l'ampleur de l'effet est de 0,60 en moyenne.

Certains des changements observés dans les scores du PISA peuvent être attribuables à des erreurs de mesure ou d'échantillonnage, ce qui entraîne un « retour à la moyenne » : les pays qui avaient des résultats élevés en 2000 étaient plus susceptibles d'avoir des taux de croissance annuelle négatifs, et ceux qui avaient des résultats faibles en 2000 étaient plus susceptibles d'avoir des taux de croissance annuelle positifs. La corrélation entre les scores des pays au PISA en 2000 et les taux de croissance annuelle est de -0,39. La Finlande, soit le pays ayant obtenu le score le plus élevé en 2000, a connu le déclin le plus prononcé : ses scores ajustés au SSÉ ont diminué de presque 2,5 % d'un écart type par année.

¹ Les estimations de la figure 1 sont basées sur l'ensemble des échantillons de données recueillis pour chaque pays lors de chaque cycle et ont été ajustées pour tenir compte des changements dans le SSÉ des échantillons. Les analyses sont basées sur un modèle de « stabilité » hiérarchique à deux niveaux adapté à chaque pays, avec le SSÉ comme covariable au niveau 1 et l'année au niveau 2 (voir Willms et Raudenbush, 1989).



Figure 1. Taux de croissance annuelle de la performance en lecture, PISA 2000 à 2015



Source: PISA 2015.

Trois prémisses sous-tendent l'approche adoptée dans le présent rapport. Premièrement, le développement des compétences en lecture des enfants doit être la priorité des systèmes de suivi de l'éducation. Il est essentiel de savoir lire pour développer de solides compétences au premier et au deuxième cycle du secondaire, pour terminer sa scolarité et pour justice sociale (Beswick, Sloat et Willms, 2008; Snow, Burns et Griffin, 1998; Willms, 2006). Deuxièmement, les recherches sur les effets de l'école et du groupe ont montré qu'il faut créer des systèmes de suivi de l'éducation qui sont informatifs. Les études nationales ou internationales de grande envergure ne doivent plus chercher les effets de l'école avec diverses mesures des facteurs associés à l'école et au groupe. Elles doivent plutôt se concentrer sur un petit nombre de facteurs, les mesurer plus en détail et les suivre longitudinalement. Troisièmement, les résultats des études internationales de grande envergure, combinés à ceux des études nationales et des petites études expérimentales contrôlées, peuvent fournir aux administrateurs de l'éducation de l'information qui les aidera à fixer des objectifs atteignables, à affecter les ressources et à évaluer les effets des politiques qui modifient une ou plusieurs caractéristiques structurelles des écoles.

Le présent rapport ne défend pas l'abandon des études internationales de grande envergure. D'ailleurs, bon nombre des exemples qui y sont cités sont basés sur des données du PISA. La majorité des pays à revenu faible et moyen n'ont pas encore participé à une évaluation internationale. Il serait avantageux pour eux de comprendre comment leurs élèves se comparent à ceux d'autres pays. Par ailleurs, les études comparatives donnent souvent aux pays la volonté politique d'investir davantage en éducation (Singer et Braun, 2018). Le présent rapport a davantage comme objectif de détourner l'attention du classement des pays sur la base de leurs résultats ou des liens de cause à effet établis à partir de données transversales.

Les exemples des sections 3 et 4 sont basés sur les données du Mexique au PISA 2015. L'intention ici n'est toutefois pas de faire l'évaluation du système d'éducation du Mexique. Backhoff, Bouzas, Hernández et García (2007) ont réalisé une évaluation détaillée des résultats du Mexique en se basant sur la structure



établie dans le rapport sur les écarts en matière d'apprentissage (Willms, 2006). Martínez et Díaz (2016) ont préparé des résultats détaillés pour le Mexique basés sur le PISA 2015.

L'objectif général du présent rapport est de proposer une structure pour l'utilisation des données de suivi afin d'orienter deux types de politiques éducatives : celles qui visent l'amélioration des résultats scolaires et la réduction des inégalités, et celles relatives aux stratégies à employer pour atteindre des objectifs éducatifs. Dans le rapport, le terme « stratégie » est employé au sens large et englobe les plans d'action; ceux-ci peuvent avoir divers objectifs, notamment la réattribution des ressources entre des écoles ou des sous-populations d'élèves, l'adoption de nouveaux programmes d'études ou de nouvelles approches pédagogiques ou la modification d'importantes caractéristiques structurelles des écoles. Quinze statistiques clés et cinq approches graphiques ont été utilisées. Les détails techniques pour l'estimation des 15 statistiques figurent à l'annexe 2, et des exemples des approches graphiques sont présentés tout au long du rapport.

La section 1 présente les grandes lignes du cadre de prospérité éducative. La section 2 traite des deux périodes de transition les plus critiques pour le développement des compétences de littératie. La section 3 aborde les politiques visant la performance des élèves. Comment peut-on utiliser les données de suivi pour répondre à cette question : « Où en sommes-nous? » Cette section traite également du repérage des élèves vulnérables et de l'établissement d'objectifs réalistes pour assurer leur développement optimal. La section 4 traite des stratégies à adopter pour atteindre des objectifs éducatifs. Elle décrit cinq types de stratégies et cherche à établir comment les données de suivi peuvent être utilisées pour déterminer quelles stratégies pourront le mieux améliorer la performance des élèves. L'objectif d'une stratégie ou d'une intervention peut être de renforcer les bases de la réussite, de réduire les inégalités dans l'affectation des ressources ou de modifier une ou plusieurs caractéristiques structurelles des écoles. Chacun de ces objectifs est abordé dans cette section. La section 5 propose une façon d'utiliser les données de suivi pour établir des objectifs et présente un exemple de système de suivi basé sur le cadre de prospérité éducative. La section se conclut par une discussion sur ce qu'implique un suivi des écoles et des systèmes scolaires basé sur le cadre de prospérité éducative.



I. Prospérité éducative : une approche du parcours de vie

S'épanouir

Un élève s'épanouit lorsqu'il développe son plein potentiel. L'objectif est de *donner donner à chaque enfant la possibilité de s'épanouir*. On dit qu'il y a prospérité éducative lorsque les familles, les communautés et les institutions publiques réussissent à assurer le bien-être scolaire, social, émotionnel, physique et spirituel des enfants. Le cadre de prospérité éducative est un outil d'évaluation dont peuvent se servir les pays, les territoires scolaires et les écoles pour suivre leur progrès vers l'atteinte de cet objectif.

La **figure 2** présente le cadre de prospérité éducative (Willms, 2018a). Le cadre comprend six stades de développement qui vont de la conception au deuxième cycle du secondaire. À chaque stade est associé un petit ensemble de résultats, les « résultats de prospérité », et des facteurs familiaux, institutionnels et communautaires qui influent sur ces résultats, les « bases de la réussite ».

Figure 2. Le cadre de prospérité éducative

	 PRENATAL	 PETITE ENFANCE 0 à 2 ans	 PRÉPRIMAIRE 3 à 5 ans	 PREMIER CYCLE DU PRIMAIRE 6 à 9 ans	 DEUXIÈME CYCLE DU PRIMAIRE ET PREMIER CYCLE DU SECONDAIRE 10 à 15 ans	 DEUXIÈME CYCLE DU SECONDAIRE 16 à 18 ans
Résultats de prospérité	<ul style="list-style-type: none"> - Grossesse en santé - Accouchement en santé 	<ul style="list-style-type: none"> - Développement langagier - Développement cognitif - Développement physique - Autorégulation 	<ul style="list-style-type: none"> - Conscience de soi et de son environnement - Compétences sociales et approches pour l'apprentissage - Développement langagier - Développement cognitif - Développement physique 	<ul style="list-style-type: none"> - Niveau de scolarité - Maîtrise de la lecture - Numératie - Santé et bien-être - Engagement 	<ul style="list-style-type: none"> - Niveau de scolarité - Réussite scolaire - Santé et bien-être - Engagement 	<ul style="list-style-type: none"> - Achèvement scolaire - Citoyen éthique - Compétences de leadership - Santé et bien-être - Compétences de communication et d'interaction
Bases de la réussite Facteurs familiaux	<ul style="list-style-type: none"> - Alimentation - Aucune exposition à des toxines - Santé physique de la mère - Santé émotionnelle de la mère 	<ul style="list-style-type: none"> - Allaitement et alimentation - Santé physique de la mère - Santé émotionnelle de la mère - Compétences parentales - Relations intrafamiliales 	<ul style="list-style-type: none"> - Compétences parentales - Relations intrafamiliales - Implication de la famille 	<ul style="list-style-type: none"> - Compétences parentales - Relations intrafamiliales - Implication de la famille 	<ul style="list-style-type: none"> - Compétences parentales - Relations intrafamiliales - Implication de la famille 	<ul style="list-style-type: none"> - Compétences parentales - Relations intrafamiliales - Implication de la famille
Facteurs institutionnels	Établissement de soins de santé : <ul style="list-style-type: none"> - Soins prénataux - Soins de santé de base 	Établissement de soins de santé : <ul style="list-style-type: none"> - Soins postnatals - Soins de santé primaires 	Établissement préscolaire : <ul style="list-style-type: none"> - Conçu pour l'enfant - Articulé autour d'objectifs - Occasions de socialiser 	École : <ul style="list-style-type: none"> - Sûre et inclusive - Enseignement de qualité des compétences langagières et de décodage - Temps d'apprentissage - Ressources matérielles 	École : <ul style="list-style-type: none"> - Sûre et inclusive - Enseignement de qualité - Temps d'apprentissage - Ressources matérielles 	École : <ul style="list-style-type: none"> - Sûre et inclusive - Enseignement de qualité - Occasions d'apprendre des compétences professionnelles et quotidiennes - Ressources matérielles
Facteurs communautaires	<ul style="list-style-type: none"> - Capital social - Ressources 	<ul style="list-style-type: none"> - Capital social - Ressources 	<ul style="list-style-type: none"> - Capital social - Ressources 	<ul style="list-style-type: none"> - Capital social - Ressources 	<ul style="list-style-type: none"> - Capital social - Ressources 	<ul style="list-style-type: none"> - Capital social - Ressources



Résultats de prospérité

Les résultats de prospérité sont considérés comme universels parce qu'ils sont essentiels au développement et à l'épanouissement de tous les enfants. Les priorités peuvent être différentes d'un pays à l'autre, surtout au dernier stade, mais les résultats de prospérité sont considérés comme souhaitables dans tous les contextes. Ces marqueurs cadrent également avec les Objectifs de développement durable (ODD) fixés par l'UNESCO (Institut de statistique de l'UNESCO, 2017).

Bases de la réussite

Les bases de la réussite sont également considérées comme universelles parce qu'un important corpus de recherches confirme qu'elles sont des conditions essentielles à la réussite à chaque stade de développement. Ces bases ont été choisies en fonction de trois critères : elles devaient être fortes, généralisées et proximales.

Un facteur « fort » a une forte corrélation avec un résultat, et la recherche reconnaît son lien *causal* avec lui. Un facteur est considéré comme *causal* s'il a été démontré qu'il précède un résultat, qu'il est corrélé à celui-ci et qu'une modification du facteur entraîne une modification du résultat (Kraemer *et. al.*, 1997). L'enseignement de qualité est considéré comme un facteur fort parce qu'il est corrélé à la réussite tout au long du parcours scolaire et parce que les stratégies pour l'améliorer entraînent une plus grande réussite (Anderson, 2004; Creemers et Kyriakides, 2006; Kyriakides, Christoforou et Charalambous, 2013; Rosenshine, 2010).

Un facteur « généralisé » est positivement corrélé à une vaste gamme de résultats et, idéalement, la recherche démontre son lien causal avec ces résultats. Par exemple, l'école « sûre et inclusive », laquelle est une base de la réussite aux trois derniers stades de développement, influe sur une vaste gamme de résultats, comme la réussite, le niveau de scolarité, l'engagement des élèves, et la santé et le bien-être.

Un facteur « proximal » a un lien direct avec un résultat; son effet ne dépend pas d'un ou de plusieurs autres facteurs. Le temps d'apprentissage a une incidence directe sur les résultats des élèves; ses effets ne dépendent pas d'autres facteurs. Le perfectionnement des enseignants n'est pas considéré comme une base de la réussite parce que ses effets dépendent de plusieurs autres facteurs. Les effets du perfectionnement se produisent que s'il donne lieu à un enseignement de meilleure qualité, à un milieu plus inclusif ou à un soutien accru de la famille et de la communauté.



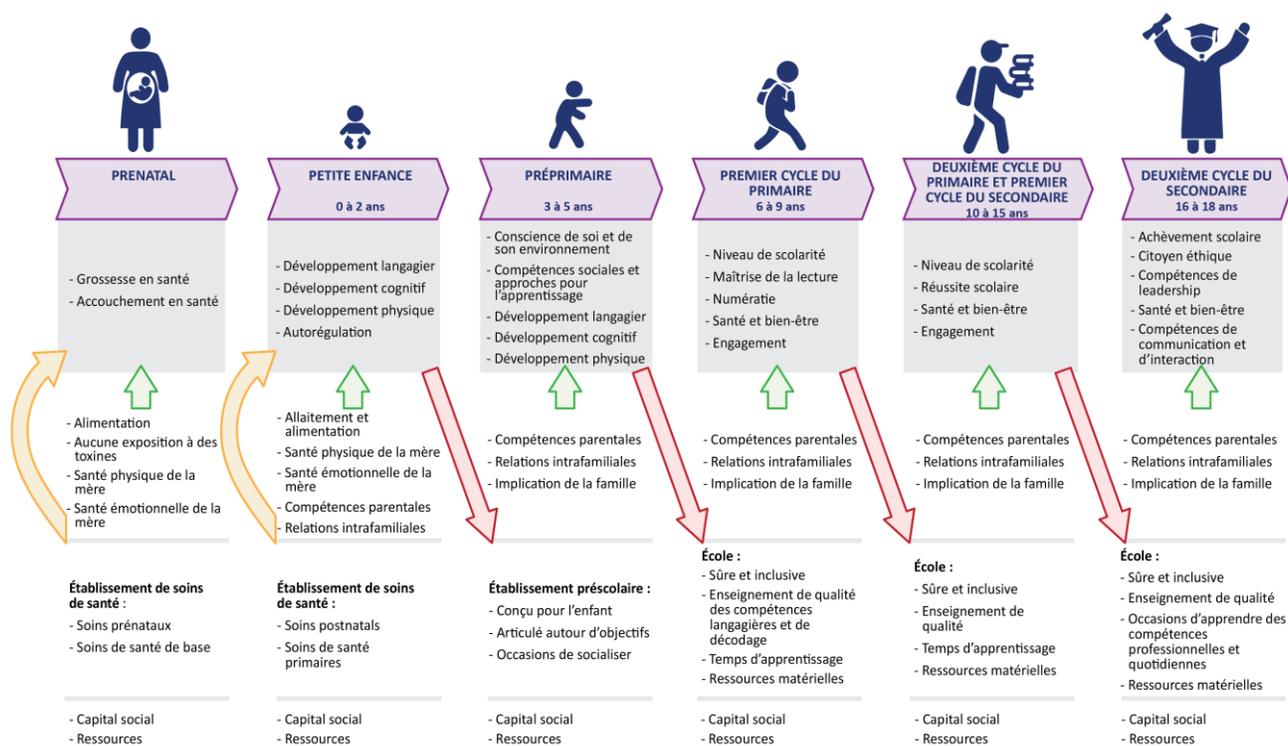
Figure 3. Quatre façons de cumuler la réussite

Flèches orange : intégration biologique

Flèches vertes : bases de la réussite

Flèches violettes : effets cumulatifs

Flèches rouges : effets de la sélection institutionnelle



Quatre façons de cumuler la réussite

Le cadre de prospérité éducative est basé sur un modèle de développement selon lequel il y a quatre façons de cumuler la réussite de la conception à la fin de l'adolescence. Ces quatre façons sont présentées à la **figure 3**.

Intégration biologique. Les résultats à la naissance sont affectés par les bases de la réussite (flèche verte) de la période prénatale (alimentation, aucune exposition à des toxines, santé physique et émotionnelle de la mère). Dans une certaine mesure, ces résultats sont intégrés biologiquement (flèche orange) par des processus épigénétiques lors desquels des signatures chimiques sont attachées aux gènes et prédisposent l'enfant à la vulnérabilité ou à la résilience (Boyce et Kobor, 2015). Les expériences vécues en bas âge interagissent avec les prédispositions génétiques et influent sur le développement du cerveau et sur d'autres systèmes neurologiques et biologiques associés au développement de l'enfant (Boyce, Sokolowski et Robinson, 2012).



À la naissance, les enfants ont des milliards de connexions neuronales nommées synapses. Pendant la petite enfance, certaines synapses sont renforcées en réponse aux stimulus de l'environnement. Les synapses qui ne sont pas utilisées sont alors élaguées afin de permettre à celles qui le sont de se renforcer davantage. Les connexions de la vision et de l'ouïe sont établies très tôt; elles sont suivies de près par celles de la maîtrise des émotions, du langage et de la cognition (Center on the Developing Child, 2007; Knudsen, 2004; Shonkoff et Phillips, 2000). Ce processus de renforcement et d'élagage des synapses (le modelage du cerveau) est plus rapide pendant certaines *périodes critiques* des deux ou trois premières années de vie (McEwen et Schmeck; 1994; Cynader et Frost, 1999; Hertzman, 1999).

Bases de la réussite. Après la naissance, le développement des enfants est soutenu par les bases de la réussite. Par exemple, de la naissance à l'âge de 2 ans, les interactions avec les parents et les gardiens sont critiques pour le développement du langage et de l'autorégulation (McClelland *et. al.*, 2010). Par conséquent, le style parental et les relations intrafamiliales sont considérés comme deux des bases de la réussite à ce stade ainsi qu'à tous les stades subséquents. Au stade préscolaire et aux trois stades scolaires qui suivent, les bases de la réussite sont les suivantes : école sûre et inclusive, enseignement de qualité, temps d'apprentissage, ressources matérielles et soutien de la famille et de la communauté.

Effets cumulatifs. Les enfants cumulent les résultats de prospérité d'un stade à l'autre. Les compétences langagières se développent tout au long de l'enfance. Certaines compétences sont toutefois des prérequis au développement d'autres compétences. Par exemple, un enfant qui a développé de solides compétences cognitives au stade préscolaire aura sans doute plus d'aisance en lecture au primaire (Leppänen *et. al.*, 2004; Nation et Snowling, 2004). Les facteurs familiaux jouent un rôle important dans le développement des résultats de prospérité. Toutefois, après l'âge de 5 ans, l'école joue aussi un rôle important qui va aller en augmentant au fil des années. Par conséquent, un enfant qui ne développe pas de solides compétences pendant ses premières années de vie risque davantage d'avoir de la difficulté à l'école.

Sélection institutionnelle. Lorsque les élèves réussissent bien à un stade de développement, leur parcours peut être altéré s'ils sont dirigés vers un groupe, une école ou un programme spécial. De nombreux systèmes scolaires dirigent les enfants qui ont de bonnes compétences en lecture et langagières vers des groupes ou des programmes qui favorisent le développement de solides bases. Ces enfants pourront davantage profiter d'interactions positives avec des pairs, d'un enseignement de meilleure qualité et d'autres privilèges qui leur permettront de développer plus rapidement leurs compétences (Willms, 2006).

Égalité, équité et accès

La Convention des Nations Unies relatives aux droits de l'enfant (1989) contient des énoncés qui reconnaissent à l'enfant le droit à un niveau de vie suffisant pour permettre son développement physique, mental, spirituel, moral et social, aux meilleurs soins de santé possibles et à une éducation de qualité, en vue d'assurer l'exercice de ce droit progressivement et sur la base de l'égalité des chances. La majorité des pays reconnaissent que leur développement socioéconomique passe par un accès universel à l'éducation de la petite enfance à l'adolescence et ont par conséquent établi des garanties constitutionnelles et juridiques pour assurer cet accès universel (Levin, 2009). Toutefois, même dans les pays les plus riches, les

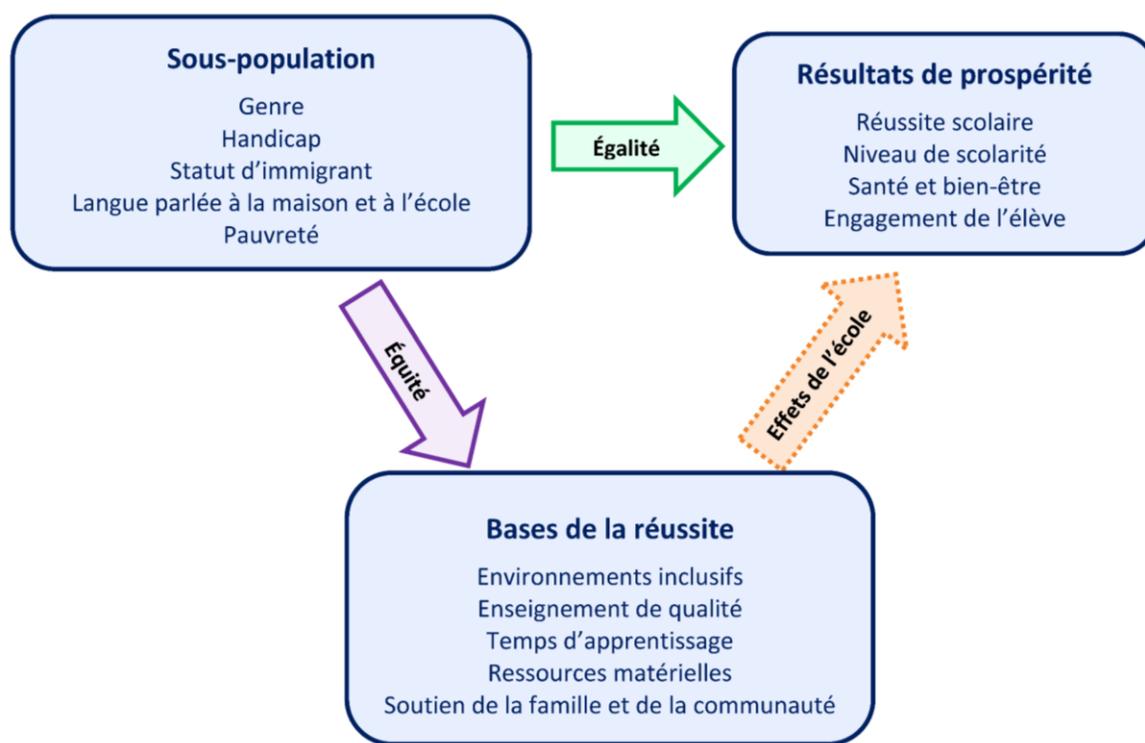


enfants issus des milieux socioéconomiques plus élevés ont davantage accès aux ressources économiques, sociales et éducatives que ceux issus des milieux socioéconomiques plus bas (Willms, 2006). Par conséquent, les leaders de l'éducation veulent de plus en plus suivre et quantifier ces inégalités dans les résultats de sous-populations, et déterminer les facteurs qui y sont associés.

Les chercheurs ont employé les termes « égalité », « équité » et « accès » de diverses manières et souvent de manière interchangeable. Le présent rapport, à l'instar de Willms, Tramonte, Duarte et Bos (2012) et de Willms (2011), fait la distinction entre l'égalité et l'équité, et voit l'accès comme un aspect de l'équité. Le modèle de prospérité éducative a été adopté par le PISA pour le développement (PISA-D), une initiative de l'OCDE dont l'objectif est d'accroître la pertinence des évaluations du PISA pour les pays à revenu faible et moyen (OCDE, 2017). Le modèle pour la prospérité éducative employé par le PISA-D est présenté à la **figure 4** pour illustrer les concepts.

Égalité. Le terme « égalité » fait référence aux différences dans la distribution des résultats scolaires de différentes sous-populations. À la figure 4, la flèche verte montre le sens du lien entre les caractéristiques démographiques (p. ex. genre ou handicap) et les résultats de prospérité. La mesure de l'égalité est relativement directe. Par exemple, la différence entre le score moyen des filles et celui des garçons à l'évaluation de lecture du PIRLS est une mesure de l'égalité. Il faut noter toutefois que la définition fait référence aux différences dans la *distribution* des résultats. En lecture, il y a une importante différence dans la distribution des scores des filles et des garçons. La prévalence de filles et de garçons qui n'atteignent pas un standard minimal est un marqueur pertinent d'égalité. Les différences dans les résultats de différents groupes socioéconomiques constituent également une mesure d'égalité. Plus loin dans le rapport, des gradients socioéconomiques sont employés pour évaluer les inégalités.

Figure 4. Égalité, équité et effets de l'école sur la prospérité éducative



Équité. L'équité est le traitement juste des personnes de différentes sous-populations. Il fait référence aux différences entre les sous-populations en ce qui a trait à l'accès à l'école et aux ressources, et aux processus qui ont une incidence sur les résultats scolaires. La flèche violette de la figure 4 montre le sens du lien entre les caractéristiques démographiques et les bases de la réussite. L'équité est difficile à mesurer parce qu'il faut tout d'abord identifier les ressources et les processus scolaires qui ont une incidence sur les résultats scolaires. À la figure 4, la flèche orange montre les effets de l'école. Ces effets sont présumés « causals ». Les études transversales de grande envergure ne permettent pas de déterminer quels sont les facteurs qui produisent des effets forts. Les bases de la réussite du cadre de prospérité éducative ont été établies à partir d'une abondante documentation.

Il est utile de faire la distinction entre égalité et équité parce que les mesures de l'égalité montrent l'ampleur des différences entre les sous-populations (c.-à-d. l'ampleur du problème), alors que les mesures de l'équité mettent en lumière ce qui doit être fait pour réduire les inégalités (c.-à-d. les politiques qui doivent être prises).

Accès. L'accès est la possibilité pour les enfants d'un territoire donné d'aller à l'école gratuitement. Selon le modèle de prospérité éducative, l'accès à l'école est un aspect des ressources matérielles et soulève donc des questions sur l'équité. Par exemple : « Les enfants qui ont un handicap ont-ils autant accès à l'école que ceux qui n'ont pas de handicap? ». Le modèle met l'accent sur l'offre scolaire. Il incombe aux gouvernements



et aux institutions scolaires de s'assurer qu'il y a des écoles dans chaque région et que celles-ci sont sûres et dotées de matériel d'apprentissage adéquat. La politique éducative ne doit pas empêcher les enfants de fréquenter l'école. À titre de base de la réussite, l'accès est considéré comme un facteur lié à l'offre.

Selon le modèle de prospérité éducative, la fréquentation scolaire et la progression d'un niveau à l'autre sont des aspects du niveau de scolarité, lequel est un résultat de prospérité. Une mesure de la fréquentation scolaire intègre des facteurs liés à la demande. Si l'offre scolaire est adéquate, si les écoles sont sûres et inclusives et dotées de ressources matérielles adéquates, si l'enseignement est de qualité, si le temps d'apprentissage est suffisant et si les écoles reçoivent le soutien des familles et de la communauté, il faut alors se demander dans quelle mesure les enfants fréquentent l'école. Ceci dépend de plusieurs facteurs culturels, sociaux, religieux, politiques et économiques.

L'énoncé 4.1 des Objectifs de développement durable des Nations Unies est le suivant : « D'ici à 2030, faire en sorte que toutes les filles et tous les garçons suivent, sur un pied d'égalité, un cycle complet d'enseignement primaire et secondaire gratuit et de qualité, qui débouche sur un apprentissage véritablement utile » (Nations Unies, 2015). Cet énoncé comprend la notion d'accès telle qu'elle est définie dans le présent rapport, à savoir une éducation gratuite (ressources matérielles), de qualité (enseignement de qualité, environnement sûr et inclusif) et menant à la réussite (soit un résultat de prospérité). La distinction entre équité et égalité est importante pour faire le suivi des progrès vers l'atteinte de cet objectif. Par exemple, on peut d'abord se demander si l'éducation offerte est équitable, puis si cette équité est liée à l'égalité de la réussite et du niveau de scolarité atteint. Le modèle CREATE de Lewin (2015) est utile parce qu'il insiste sur l'importance d'avoir des effectifs qui réussissent et qui atteignent un bon niveau de scolarité, de même que sur les « zones d'exclusion » associées au décrochage scolaire.



Sous-populations. Les sous-populations d'intérêt peuvent ne pas être les mêmes d'un pays à l'autre et selon le résultat de prospérité ou le stade de développement. Par exemple, les objectifs du Programme de développement durable à l'horizon 2030 (Institut de statistique de l'UNESCO, 2017) mettent l'accent sur les inégalités et sur les iniquités associées au genre. Selon les pays qui participent au PISA pour le développement, les facteurs énumérés à la figure 4 (genre, handicap, statut d'immigrant, langue parlée à la maison et à l'école, et pauvreté) sont les plus importants dans leurs contextes.

Le rôle du SSÉ

Le SSÉ et la pauvreté jouent un rôle important tout au long de l'enfance et sont liés aux bases familiales et communautaires à tous les stades de développement. Le SSÉ et la pauvreté ont également une incidence sur l'accès des enfants à des bases institutionnelles solides, surtout après l'âge de 5 ans, quand la sélection institutionnelle est plus marquée. La **figure 5** montre le rapport entre le pourcentage d'élèves ayant obtenu un score de niveau 2 ou inférieur au PISA 2015 et le revenu national brut par habitant. Le graphique montre clairement que la vulnérabilité augmente de façon prononcée lorsque le revenu national brut par habitant est inférieur à 30 000 \$.

L'approche du parcours de vie adoptée par le cadre de prospérité éducative suggère que les enfants deviendraient vulnérables tôt dans leur vie, et plusieurs le seraient déjà avant même d'entrer à l'école. Ces enfants démontrent sans doute déjà un potentiel pour devenir des lecteurs accomplis lorsqu'ils font leur entrée à l'école primaire. Dans la majorité des pays à revenu faible et moyen, plus de 60 % des élèves ont de faibles compétences en lecture à l'âge de 15 ans. Dans les pays à revenu élevé, entre 30 % et 50 % des élèves de 15 ans ont de faibles compétences en lecture. Nous ne savons pas dans quelle mesure le potentiel de développer de bonnes compétences en littératie est biologiquement intégré au stade prénatal ou au stade de la petite enfance et au stade préprimaire, mais les données laissent croire de plus en plus que les premières années de vie joueraient un rôle dominant dans le développement de ces compétences. Le PISA pour le développement comprend une mesure détaillée du SSÉ qui englobe l'extrémité inférieure de l'échelle de SSÉ afin de détecter les enfants vivant dans une pauvreté grave ou extrême.

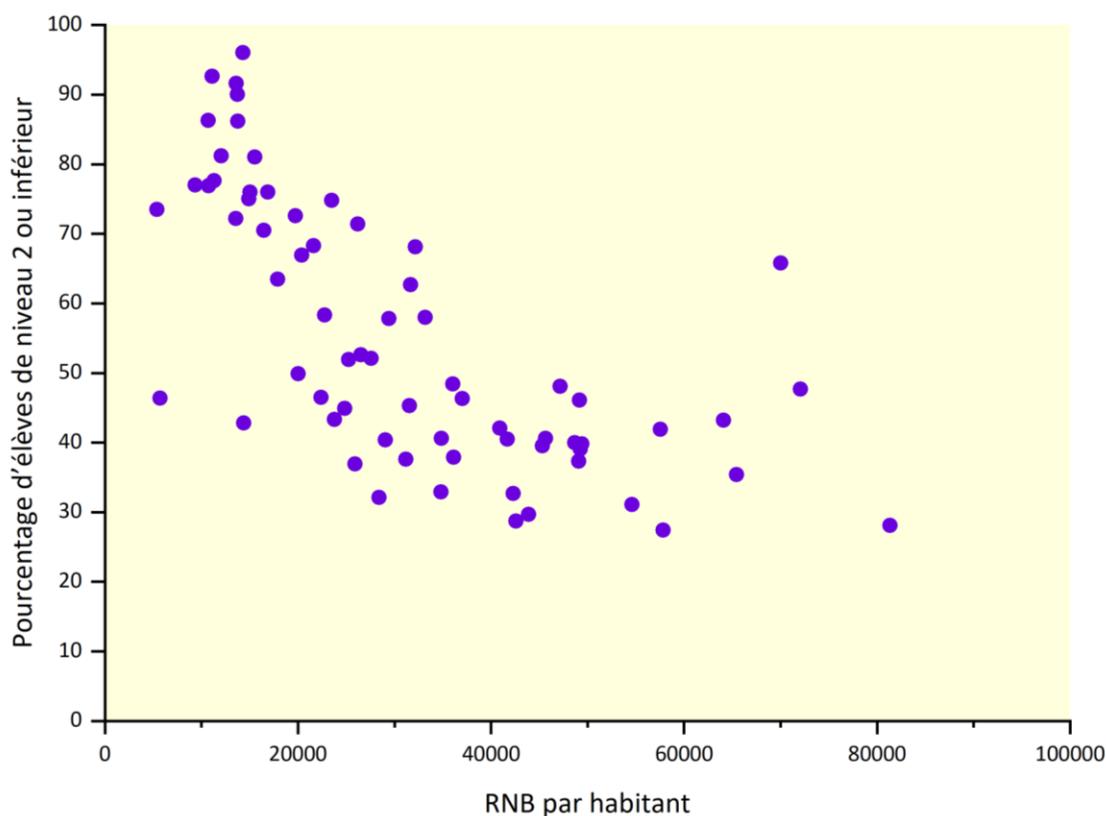
On ne peut toutefois pas présumer que les dés sont jetés pour les enfants au stade préprimaire. En effet, dans tous les pays, les enfants peuvent développer des compétences de prélecture lorsqu'ils commencent l'école primaire. Un système scolaire fort soutenu par les familles et par la communauté a le pouvoir d'influer sur le développement subséquent de tous les enfants. La majorité des enfants peuvent apprendre à lire s'ils reçoivent un enseignement de qualité dans un environnement inclusif et si le temps d'apprentissage est suffisant.

Tous les stades du cadre de prospérité éducative sont sans doute critiques pour le développement des compétences de littératie. Toutefois, deux périodes de transition sont particulièrement importantes lorsqu'il s'agit de faire le suivi d'un système scolaire dans le but d'orienter la politique éducative. La première période de transition critique est l'entrée à l'école primaire. Lorsqu'il entre au primaire – autour de l'âge de 6 ans dans la majorité des systèmes – l'enfant passe de la maison, de la garderie, de la maternelle ou d'un programme d'éducation de la petite enfance à une salle de classe de première année, laquelle est un milieu



d'apprentissage structuré suivant un programme plus officiel. La deuxième période critique est celle qui marque le passage du premier cycle au deuxième cycle du primaire, lorsque l'enfant n'apprend plus seulement à lire, mais lit pour apprendre. Ces deux périodes de transition sont abordées plus en détail dans la prochaine section.

Figure 5. Pourcentage des élèves ayant des compétences en lecture de niveau 2 ou inférieur à l'âge de 15 ans par rapport au revenu national brut (RNB) (méthode Atlas en dollars américains)



Sources : PISA 2015 et Banque mondiale, 2018a.



II. Deux périodes de transition critiques pour le développement des compétences en littératie

Entrée à l'école primaire

Les gouvernements reconnaissent de plus en plus que les investissements dans les programmes d'éducation de la petite enfance sont très bénéfiques pour le développement des enfants (Knudsen, Heckman, Cameron et Shonkoff, 2006). L'entrée à l'école primaire étant une période de transition critique, il est utile de mesurer le développement des enfants à ce moment, et ce, pour deux raisons. Cette mesure est un indicateur tardif de la réussite d'une société à préparer les enfants pour l'école. Elle est également un indicateur précurseur pour repérer les enfants qui pourraient avoir besoin de soutien additionnel pendant leurs premières années à l'école primaire.

Le terme « maturité scolaire » est souvent employé pour indiquer dans quelle mesure un enfant est prêt à évoluer dans un environnement plus structuré (UNICEF, 2012). Une définition plus large du terme inclut les compétences de l'enfant et le rôle de la famille et de la communauté (Rhode Island Kids Count, 2005) :

Familles prêtes + Communautés prêtes + Services prêts + Écoles prêtes = Enfants prêts pour l'école

Le modèle élaboré pour mesurer la maturité scolaire par le U.S. National Education Goals Panel comprend cinq domaines : bien-être physique et développement moteur, développement social et émotionnel, méthodes d'apprentissage, développement du langage et cognition, et connaissances générales (Barnett, Ayers et Francis, 2015; Kagan, Moore et Bredekamp, 1995; National Early Literacy Panel, 2008).

Le lien entre les compétences cognitives et langagières à l'âge de 5 ans et les compétences en lecture à l'âge de 8 ou 9 ans est bien établi (Deary, Strand, Smith et Fernandes, 2007; Duncan *et al.*, 2007; McClelland, Morrison et Holmes, 2000; Raver *et al.*, 2011; Rose, 2006). Les compétences en langue parlée et les compétences cognitives sont particulièrement importantes (Scarborough, 2001). L'attention et l'autorégulation sont également essentielles, et semblent être plus importantes que les troubles externalisés du comportement (Trezesniewski, Moffit, Caspi, Taylor et Maughan, 2006).

L'Évaluation de la petite enfance (ÉPE) (*Early Years Evaluation* (EYE) en anglais) est employée pour évaluer le développement des enfants de 3 à 6 ans, alors qu'ils se préparent à faire la transition vers l'éducation formelle (The Learning Bar, 2011). Cette évaluation comprend des mesures des cinq domaines établis par le *National Education Goals Panel*. Elle comprend deux outils complémentaires qui aident les éducateurs à suivre le développement global des enfants qui s'appêtent à faire la transition vers l'école. L'ÉPE – Appréciation directe (ÉPE-AD) est une évaluation interactive basée sur le jeu menée par un évaluateur formé. L'ÉPE – Appréciation de l'enseignante (ÉPE-AE) fournit un cadre systématique dont peuvent se servir les enseignantes pour structurer leurs observations et évaluations informelles. Elle est généralement utilisée par les enseignantes de maternelle pour obtenir de l'information utile et pertinente qui orientera leur enseignement.



L'ÉPE-AE a été utilisée lors d'une évaluation menée à l'échelle nationale en Uruguay en 2017. La mise en œuvre de cette évaluation a nécessité trois ans de travail et la réalisation d'une étude pilote. Les buts de cette étude pilote étaient les suivants : examiner les propriétés psychométriques de l'évaluation, adapter l'évaluation au contexte culturel, en consultation avec des enseignants et des experts des programmes, faire participer les enseignants au développement du processus de mise en œuvre et les encourager à en faire une meilleure pratique, faire réviser l'évaluation par le gouvernement pour s'assurer que son contenu était pertinent pour des enfants de maternelle et qu'elle contribuait à l'amélioration des pratiques pédagogiques (Lopez, 2016; Willms, 2018b).

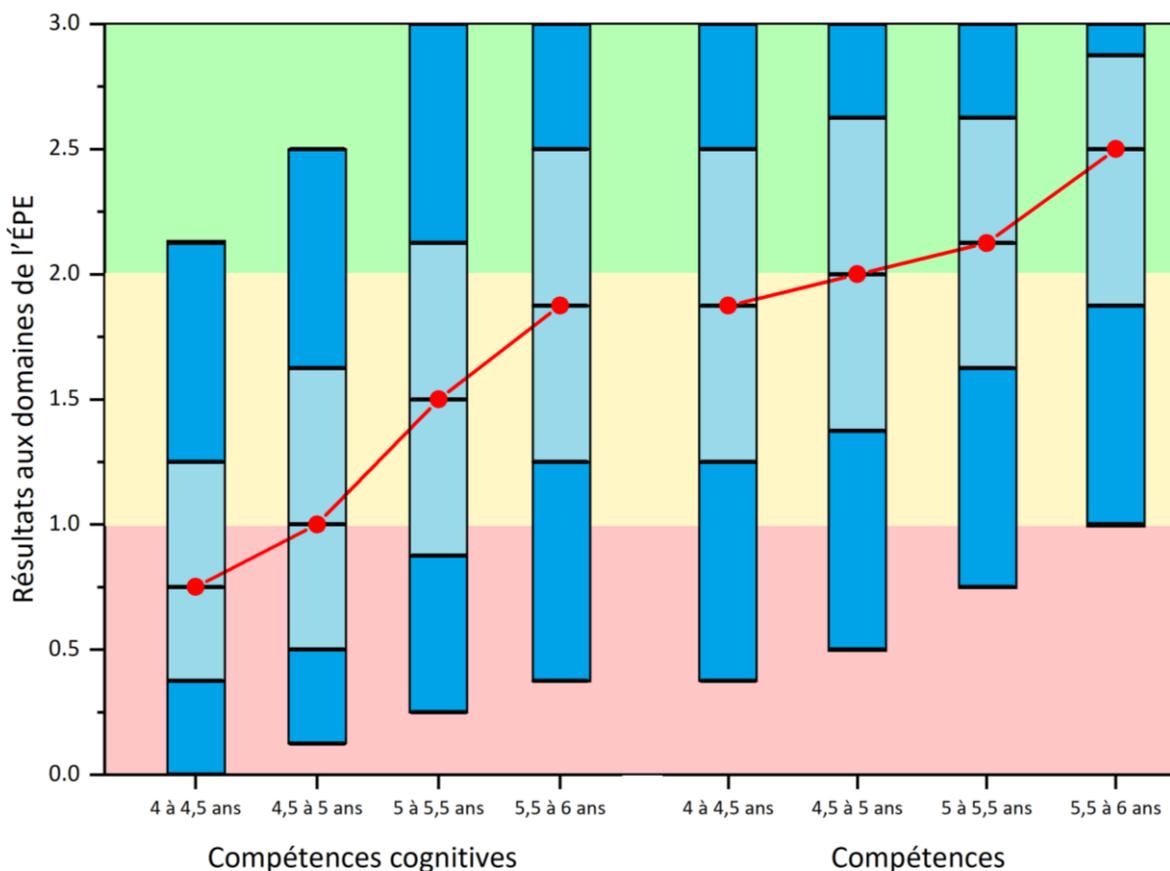
La **figure 6** montre la distribution des scores pour l'ensemble des enfants de 4 et 5 ans en Uruguay, en 2017. L'ÉPE-AE est un test fondé sur des critères qui permet d'établir si les enfants maîtrisent les compétences nécessaires pour réussir à l'école primaire. Les scores vont de zéro à trois. Les enfants qui ont obtenu un score de deux ou plus sont au niveau 1 (en vert). On considère que ces enfants ont réussi les tâches liées au domaine de développement. Les enfants au niveau 2 (score d'un ou plus, mais inférieur à deux; en jaune) ont eu certaines difficultés à réaliser les tâches, tandis que ceux au niveau 3 (score inférieur à un; en rouge) ont eu des difficultés importantes. Les diagrammes de quartiles montrent la distribution des scores pour chaque groupe d'âge (chacun correspond à une demi-année). Pour chaque diagramme, la section inférieure comprend les scores du 5^e au 25^e percentile; la section suivante les scores du 25^e percentile à la médiane; la section au-dessus les scores de la médiane au 75^e percentile; et la section supérieure les scores du 75^e au 95^e percentile.

Les résultats montrent que les compétences cognitives de la majorité des enfants de 4 à 4,5 ans sont inférieures à leurs compétences langagières. Les compétences cognitives se développent toutefois à un rythme plus rapide. Ces résultats concordent avec les résultats des recherches récentes sur le développement du cerveau. Selon ces recherches, les synapses des compétences langagières atteignent leur apogée avant les synapses des compétences cognitives, et les synapses des compétences langagières sont plus stables à l'âge de 5 ans que les synapses des compétences cognitives, lesquelles ne se stabilisent qu'à l'âge de 15 ans (Bhattacharjee, 2015).

Au point de vue des compétences cognitives, plus de la moitié des enfants de 5 ans sont considérés comme vulnérables parce qu'ils ont eu certaines difficultés ou des difficultés importantes. Chez les enfants de 5 à 5,5 ans, 38 % ont eu certaines difficultés et 33 % des difficultés importantes. Chez les enfants de 5,5 à 6 ans, ces pourcentages étaient respectivement de 34 % et de 49 %. Les enfants de 5 ans qui sont dans le quartile inférieur ont des compétences comparables à celles des enfants de 4 à 4,5 ans. Autrement dit, le développement de leurs compétences cognitives est en retard de plus d'une année par rapport à leurs pairs. La prochaine sous-section aborde le rôle des compétences cognitives dans le développement des compétences de décodage, lesquelles sont déterminantes pour l'apprentissage de la lecture.



Figure 6. Scores pour le domaine des compétences cognitives et langagières à l'Évaluation de la petite enfance, Uruguay, 2017



Source : *The Learning Bar*, 2017.

De « apprendre à lire » à « lire pour apprendre »

La transition de « apprendre à lire » à « lire pour apprendre » se fait autour de 8 ou 9 ans chez la majorité des enfants et est essentielle à la réussite scolaire, à l'atteinte d'un niveau de scolarité élevé et au bien-être (Snow, Burns et Griffin, 1998). À l'école primaire, les enfants apprennent des contenus et acquièrent diverses compétences. L'apprentissage de la lecture est toutefois central. Lorsque les enfants font la transition vers le deuxième cycle du primaire, on présume qu'ils lisent avec aisance et comprennent le contenu des matières enseignées (p. ex. santé, sciences humaines, sciences). On s'attend à ce que les enfants lisent pour apprendre – qu'ils apprennent les termes propres aux matières et qu'ils les utilisent pour exercer une pensée critique, résoudre des problèmes et acquérir de nouvelles connaissances. Plus les enfants progressent vers le premier cycle du secondaire, plus ils doivent avoir de solides compétences en lecture. Ceux qui n'ont pas les compétences fondamentales en lecture accusent des retards de plus en plus importants. Selon des études de longue date, la majorité des élèves qui ont de la difficulté en lecture à la fin du primaire continuent



d'éprouver des difficultés au premier et au deuxième cycle du secondaire (Francis *et. al.*, 1996; Warwick, 2005).

L'apprentissage de la lecture est un processus complexe qui fait intervenir la conscience phonologique (la capacité d'entendre et de manier les phonèmes d'une langue), la compréhension du principe alphabétique (comprendre qu'un graphème correspond à un phonème) et la fluidité ou la reconnaissance des mots dans un texte (Good, Simmons et Kame'enui, 2001; Storch et Whitehurst, 2002). Selon le « modèle simple de la lecture » (*Simple View of Reading* en anglais) de Gough et Tunmer (1986), l'apprentissage de la lecture dépend de l'acquisition de deux variables : le décodage et la compréhension linguistique. Le décodage est la capacité à reconnaître des mots familiers et non familiers. La capacité d'extraire correctement le code phonologique d'un mot écrit contribue non seulement au développement des compétences en lecture chez l'enfant, mais également au développement de son vocabulaire et de sa compréhension de l'écrit (Verhoeven, van Leeuwe et Vermeer, 2011). La compréhension linguistique est la capacité de comprendre et d'interpréter la langue orale et écrite à l'intérieur de phrases ou de discours. Pour qu'un enfant devienne un lecteur accompli, il doit non seulement pouvoir décoder avec précision les mots écrits, mais il doit également comprendre le sens des mots seuls et à l'intérieur d'énoncés, de phrases et de paragraphes.

La capacité à décoder des mots est le « filtre critique » à l'école primaire. Avant qu'un enfant puisse comprendre ce qu'il lit, il doit reconnaître les mots correctement et efficacement, puis retenir l'information contenue dans un énoncé ou une phrase dans sa mémoire de travail (Perfetti, Landi et Oakhill, 2005; Vellutino et Scanlon, 1987). Le développement des compétences de décodage renforce le développement des compétences de compréhension. Toutefois, la majorité des enfants qui éprouvent des difficultés en lecture ont des difficultés de décodage (Storch et Whitehurst, 2002; Verhoeven, van Leeuwe et Vermeer, 2011). Les enseignants efficaces développent une gamme de stratégies pour développer la conscience phonémique et la reconnaissance des mots chez leurs élèves. Ils obtiennent de meilleurs résultats lorsqu'ils enseignent le décodage en faisant lire à leurs élèves des textes en lien avec leur monde (Torgesen, Otaiba et Grek, 2005).

Confident Learners est un programme dont l'objectif est d'améliorer les compétences en littératie des enfants autochtones pendant le primaire (The Learning Bar Inc., 2016). Il a été élaboré pour les éducateurs autochtones du Canada et en collaboration avec eux, et avec le soutien d'un cercle autochtone d'experts. Le programme prend la forme d'un « parcours » pour l'enseignement et l'évaluation basé sur le modèle simple de la lecture. Il offre aux enseignants un programme de formation pour les aider à mieux comprendre comment se fait l'apprentissage de la littératie, et pour leur faire découvrir des stratégies efficaces pour l'enseignement du décodage et des compétences langagières. Le programme comprend 20 modules pour les compétences de décodage et 20 modules pour les compétences langagières. Chaque module comprend 14 objectifs basés sur des compétences. Chaque module propose une ou plusieurs activités d'apprentissage amusantes, intéressantes et culturellement pertinentes en lien avec les objectifs. Les enseignants évaluent régulièrement les élèves à mesure qu'ils progressent d'un module à l'autre afin de déterminer s'ils maîtrisent les objectifs.



Lors du stade initial de développement du programme *Confident Learners*, trois experts en littératie ont consulté les programmes d'études de toutes les provinces du Canada, de certains États des États-Unis, de l'Australie et du Royaume-Uni. Ils ont ensuite élaboré un modèle décrivant en détail les compétences qui sont nécessaires pour décoder les mots et développer les compétences langagières essentielles à l'apprentissage de la lecture à l'école primaire. Le modèle a répertorié plus de 300 compétences de décodage et 300 compétences langagières pour la période de la maternelle à la 3^e année. Constat inattendu : plus de la moitié des compétences de décodage essentielles sont enseignées dans la plupart des programmes du début de la 1^{re} année jusqu'à la première moitié de la 2^e année. L'équipe de recherche a décrit ce phénomène comme la « montagne de la lecture ».

Les enfants qui amorcent le primaire sans ces compétences de prélecture essentielles ne peuvent pas franchir cette montagne. Les résultats de l'Évaluation de la petite enfance montrent que, dans plusieurs territoires et dans certaines sous-populations, environ entre le quart et la moitié des enfants entrent à l'école primaire avec des compétences cognitives et langagières qui ont une année de retard ou plus sur celles de leurs pairs. Dans un programme où les enfants doivent progresser d'une année à l'autre, ces enfants sont incapables de maîtriser les compétences requises en 1^{re} année pour passer à l'année suivante. Les trois solutions le plus souvent préconisées dans de tels cas sont les suivantes : faire redoubler ces enfants, les placer dans des groupes pour enfants à besoins particuliers ou les faire passer malgré tout au niveau suivant, en espérant qu'ils réussiront à rattraper leur retard. Aucune de ces stratégies n'est efficace pour augmenter le nombre d'enfants qui réussissent à faire la transition de « apprendre à lire » à « lire pour apprendre ».

Les résultats au programme national d'évaluation de la littératie et de la numératie de l'Australie (*National Assessment Program of Literacy and Numeracy*) montrent clairement le phénomène de la « montagne de la lecture » et la lente amélioration des compétences en lecture après l'école primaire (NAPLAN; Australian Curriculum, Assessment and Reporting Authority, 2017). La **figure 7** montre la distribution des scores en lecture de la 3^e année (âge moyen de 8 ans et 3 mois) à la 9^e année (âge moyen de 14 ans et 7 mois) des élèves autochtones et non autochtones. Les diagrammes de quartiles montrent la distribution des scores pour chaque année. La section inférieure de chaque diagramme comprend les scores du 5^e au 20^e percentile; la section suivante les scores du 20^e percentile à la moyenne; la section au-dessus les scores de la moyenne au 80^e percentile; et la section supérieure les scores du 80^e au 95^e percentile.

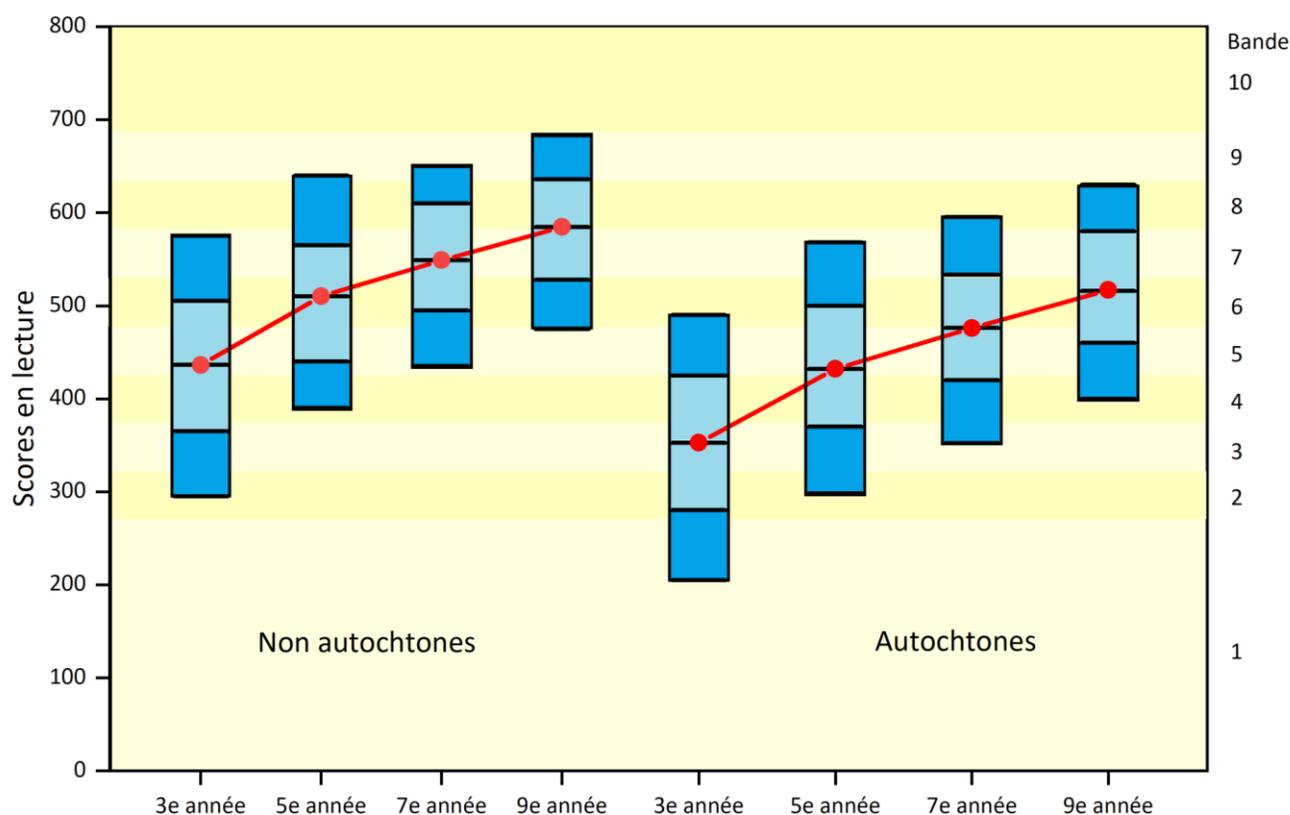
Ces résultats sont pertinents à trois égards pour le développement des compétences en lecture des élèves. Premièrement, le taux de croissance est lent entre la 3^e et la 5^e année, et encore plus lent après la 5^e année. Le score moyen pour les élèves non autochtones est de 436,6 (ÉT = 84,1) en 3^e année et de 510,2 (ÉT = 74,4) en 5^e année. Le taux de croissance moyen par année, exprimé en ampleur de l'effet, est de 0,44 par année. Le taux de croissance moyen par année est de 0,26 (moyenne = 548,9, ÉT = 66,7) entre la 5^e et la 7^e année, et de 0,24 (moyenne = 584,6, ÉT = 63,6.7) entre la 7^e et la 9^e année.

Deuxièmement, les taux de croissance pour les élèves autochtones sont comparables : 0,43 par année entre la 3^e (moyenne = 352,8, ÉT = 92,6) et la 5^e année (moyenne = 432,1, ÉT = 81,3); 0,27 par année entre la 5^e et



la 7^e année (moyenne = 476,0, ÉT = 75,0); 0,27 par année entre la 7^e et la 9^e année (moyenne = 516,9, ÉT = 73,0). La principale différence entre les résultats des élèves autochtones et ceux des élèves non autochtones s'observe en 3^e année. Cet écart dans les résultats peut être attribuable au vécu différent des enfants des deux groupes avant leur entrée à l'école ou au cours des premières années du primaire. Il n'existe aucune donnée qui pourrait montrer si cet écart était déjà présent à la maternelle ou si celui-ci s'est creusé entre la maternelle et la 3^e année. La compréhension de cette relation et de l'offre scolaire au primaire a d'importantes implications pour le suivi du système d'éducation.

Figure 7. Scores en lecture en Australie, NAPLAN 2017



Source : ACARA, 2017.

Les élèves du quintile inférieur en 9^e année ont des compétences comparables à celles des élèves autour de la médiane en 5^e année. Cette situation s'observe autant chez les élèves autochtones que chez les non-autochtones. Si une intervention avait été faite auprès des élèves du quintile inférieur en 5^e année en vue d'augmenter leurs scores d'un quart d'un écart type par année, ceux-ci auraient eu en 9^e année des scores qui se seraient situés dans le milieu de la distribution.

Effets du niveau scolaire dans le PISA. Le taux de croissance relativement faible des scores en lecture après la 9^e année observé en Australie – mais aussi de manière générale dans les pays de l'OCDE – est également observé dans les scores du PISA. Les élèves qui participent au PISA ont 15 ans. Dans la majorité



des pays, il y a des élèves de 15 ans dans au moins deux années scolaires, en excluant ceux qui ont redoublé. Les élèves qui sont une année sous l'année modale le sont en raison de leur date de naissance. Ils sont donc représentatifs des élèves de 15 ans qui progressent comme prévu. Par exemple, en Australie, parmi les élèves de 15 ans qui n'avaient pas redoublé, 9 % étaient en 9^e année, 76 % en 10^e année (année modale) et 15 % en 11^e année. Par conséquent, il est possible d'estimer un « effet du niveau scolaire » sur les scores en lecture par échelon d'une année (d'une année sous l'année modale à l'année modale et de l'année modale à une année au-dessus de l'année modale).

Pour tous les pays de l'OCDE, l'effet du niveau scolaire est de 55 points entre l'année modale et l'année sous l'année modale, et de 33 points entre l'année modale et l'année au-dessus de l'année modale. L'effet global du niveau scolaire est de 43 points. Pour les pays non membres de l'OCDE, les effets du niveau scolaire sur les élèves du même âge sont les suivants : 41 points entre l'année modale et l'année sous l'année modale, et 24 points entre l'année modale et l'année au-dessus de l'année modale, pour un effet global du niveau scolaire de 29 points. L'effet global du niveau scolaire pour tous les pays participants est de 36 points.²

Les points du PISA traduisent directement l'ampleur de l'effet, 100 points représentant une ampleur de l'effet de 1,0. Par ailleurs, une différence de 43 points représente une année scolaire, ce qui représente 4,3 points par mois lorsque l'année scolaire est de 10 mois.

Apprendre à lire dans d'autres langues que l'anglais. La vitesse à laquelle les enfants deviennent habiles à décoder les mots dépend de la transparence orthographique de la langue. L'anglais et le français ont une faible transparence orthographique (elles sont opaques), c'est-à-dire que la correspondance entre les graphèmes et les phonèmes est moins grande que celle des langues dites transparentes, comme le finnois, l'italien et l'espagnol. Il faut plus de temps pour apprendre à lire une langue opaque, comme l'anglais, qu'une langue transparente, comme l'espagnol (Caravolas *et. al.*, 2013). La capacité à décoder les mots est un prérequis pour lire toutes les langues alphabétiques, mais dans le cas des langues transparentes, la compréhension linguistique joue un rôle plus important dans le développement précoce de la lecture. Des études réalisées sur le développement précoce de la lecture en espagnol (sur la base du modèle simple de la lecture) confirment l'importance de la compréhension linguistique par rapport au décodage (Polo, Araujo et Salceda, 2017; Ripoll, Aguado et Castilla-Earls, 2014).

² Cette analyse a été réalisée grâce à la technique de modèles linéaires hiérarchiques (MLH), où les élèves sont emboîtés dans les pays. Par conséquent, l'effet estimé est l'effet moyen pour tous les pays. L'analyse a révélé que l'effet du niveau scolaire est très différent d'un pays à l'autre : il est de moins de 5 points dans quatre pays et de plus de 65 points dans huit pays. L'effet du niveau scolaire dans les scores du PISA de l'Australie est de 16 points entre la 9^e et la 10^e année, et de 12 points entre la 10^e et la 11^e année. Les scores du NAPLAN, dont il a été question précédemment, montrent eux aussi un ralentissement des taux de croissance après la 5^e année. (Un MLH a également été réalisé en prenant l'âge des élèves comme variable. Les résultats ont révélé que l'âge a un effet négligeable : les élèves qui étaient jeunes par rapport à leur cohorte ont obtenu moins d'un point de moins que ceux qui étaient vieux par rapport à leur cohorte.)



Ces résultats ont une importante incidence pour les interventions en lecture conçues pour les pays à revenu faible et moyen. Par exemple, dans de nombreux pays d'Afrique, les enfants apprennent à lire une langue nationale officielle qui peut ou ne pas être leur langue maternelle ou apprennent à lire leur langue maternelle au primaire, puis continuent ensuite dans une langue nationale. Si la première langue d'enseignement est la langue maternelle de l'enfant, une intervention en lecture comme celle du programme *Confident Learners* viendra peut-être développer davantage les compétences de décodage que la compréhension linguistique au primaire. La hauteur de la « montagne de la lecture » ne sera pas la même selon le niveau de transparence orthographique de la langue. Si la langue d'enseignement n'est pas la langue maternelle de l'enfant, l'intervention doit porter davantage sur le développement linguistique parce que celui-ci risque moins d'être renforcé à la maison et dans la communauté. L'importance relative à accorder aux compétences de décodage par rapport aux compétences de compréhension linguistique est un élément central des politiques d'éducation en Afrique du Sud. En 1994, quand le régime d'apartheid a pris fin, la constitution du pays a accordé le statut de langues officielles à neuf langues autochtones, aux côtés de l'afrikans et de l'anglais (Heugh, 2013). Dans la majorité des cas, les enfants reçoivent un enseignement dans leur langue maternelle jusqu'en 3^e année, puis passent ensuite à l'anglais (Manyike, 2012). Ceci peut être un problème pour la majorité des élèves qui n'ont pas acquis les compétences de décodage nécessaires pour la lecture de l'anglais.

La pauvreté est également un important facteur à prendre en considération avant de déterminer les interventions les plus appropriées pour développer les compétences de prélecture au stade préprimaire et pour enseigner la lecture au début du primaire. Le lien entre la pauvreté et les compétences de prélecture des enfants a été bien établi (Banque mondiale (2018b)). Les enfants qui vivent dans la pauvreté risquent davantage de souffrir de carence nutritive, de contracter des maladies infectieuses et de vivre dans des environnements toxiques pendant leurs deux premiers stades de développement (de la conception à l'âge de deux ans). Par ailleurs, les conséquences de ces facteurs de risque peuvent être intégrées biologiquement chez ces enfants (Black *et al.*, 2008). Les enfants pauvres risquent également de recevoir moins de stimulation, d'encadrement et de soutien pendant leurs premières années de vie (Black *et al.*, 2017). Les données pour l'Uruguay de l'Évaluation de la petite enfance, dont il a été question plus tôt, montrent que plus du quart des enfants qui entrent en 1^{re} année ont plus d'une année de retard sur leurs pairs en ce qui a trait au développement de leurs compétences cognitives. Le pourcentage d'enfants vulnérables est plus grand dans les régions où le pourcentage d'enfants vivant dans la pauvreté est plus élevé. S'il fallait élaborer une intervention en lecture pour l'Uruguay, il faudrait accorder davantage d'importance au décodage dans les régions plus pauvres et à la compréhension linguistique dans les régions plus riches.



III. Politiques sur les résultats des élèves et l'égalité

L'objectif général d'une politique éducative est de faire en sorte que tous les enfants développent leur plein potentiel et *s'épanouissent*. Une politique éducative établit un plan d'action pour fixer des objectifs et déterminer les stratégies qui permettront de les atteindre. Les données de suivi peuvent orienter les politiques parce qu'elles fournissent de l'information sur la performance du système scolaire. La principale question à se poser est la suivante : « Où en sommes-nous? ». En général, il y a trois façons de donner un sens aux indicateurs sociaux : en les comparant à une norme (p. ex. une moyenne nationale ou internationale), en comparant ceux de plusieurs territoires (p. ex. en comparant les écoles d'une province ou d'un État) et en étudiant les tendances dans le temps. Voici des exemples de questions qui peuvent être posées : « Comment se compare-t-on aux normes internationales? », « Comment se compare-t-on aux autres pays qui ont des ressources similaires? », « Nos résultats s'améliorent-ils au fil du temps? ». Les réponses à ces questions entraînent inévitablement d'autres questions : « Qui sont les enfants vulnérables? Y a-t-il des inégalités dans les résultats? », « Où sont les enfants les plus vulnérables? ». Les réponses à ces questions permettent aux administrateurs d'établir des objectifs réalistes et mesurables pour améliorer les résultats des élèves et réduire les inégalités. La présente section porte sur les résultats de prospérité. Son objectif est d'établir quels indicateurs sont les plus utiles et quels analyses et outils peuvent transmettre de l'information sur les résultats des élèves et l'égalité, et orienter les politiques. La section 5 traite des façons d'utiliser cette information pour fixer des objectifs.

Où en sommes-nous?

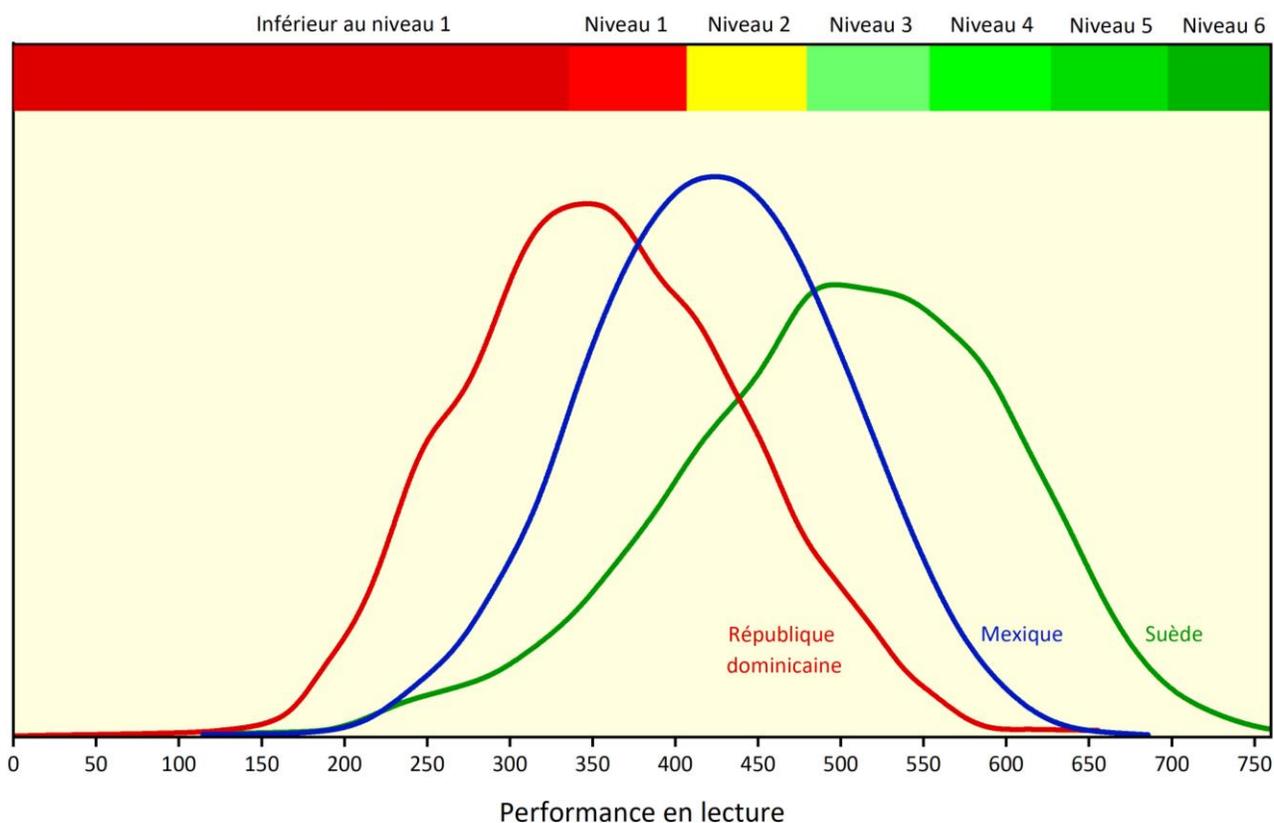
Se concentrer sur les résultats. Le niveau de scolarité atteint est le meilleur point de départ. Pour les pays à revenu faible et moyen, il faut d'abord se demander si les enfants reçoivent une éducation préprimaire ou amorcent l'éducation primaire. La majorité des cohortes d'élèves poursuivent généralement plus longtemps leurs études que celles qui les ont précédées (Banque mondiale, 2018, b). Toutefois, certains enfants ne fréquenteront jamais l'école, dans certains cas parce qu'ils habitent des régions éloignées ou encore parce qu'ils ont un handicap. Dans de nombreux pays à revenu faible, le taux de scolarisation des filles est beaucoup plus bas que celui des garçons (Banque mondiale, 2018b). Deux marqueurs clés sont donc tout simplement le pourcentage d'élèves scolarisés au préprimaire et au primaire. À ce stade, les mesures des compétences de prélecture des enfants revêtent également une importance parce qu'elles sont des indicateurs tardifs du développement depuis la conception et des indicateurs précurseurs des ressources qui seront nécessaires pendant les premières années du primaire. Avec ces indicateurs en main, on peut ensuite se demander quelles sont les sous-populations vulnérables et où habitent les enfants vulnérables.

Le pourcentage d'enfants qui font la transition du premier au deuxième cycle du primaire est un autre marqueur clé du niveau de scolarité atteint. Il est également souhaitable d'avoir des mesures des compétences en littératie et en numératie, de la santé et du bien-être, et de l'engagement des élèves au moment de cette transition. Il est essentiel d'avoir un indicateur du pourcentage d'élèves qui font cette transition. Sans cet indicateur, ceux sur la littératie et la numératie n'ont que peu de valeur parce qu'ils risquent de n'être représentatifs que d'un groupe sélectionné. Le même principe s'applique pour les



transitions du deuxième cycle du primaire au premier cycle du secondaire et du premier au deuxième cycle du secondaire. Les scores aux évaluations internationales, comme le PIRLS, la TIMSS ou le PISA, sont peu utiles à des fins de comparaison, à moins d'être comparés à des indicateurs du pourcentage d'enfants qui ne fréquentent pas l'école.

Figure 8. Distribution de la performance en lecture à l'âge de 15 ans en République dominicaine, au Mexique et en Suède



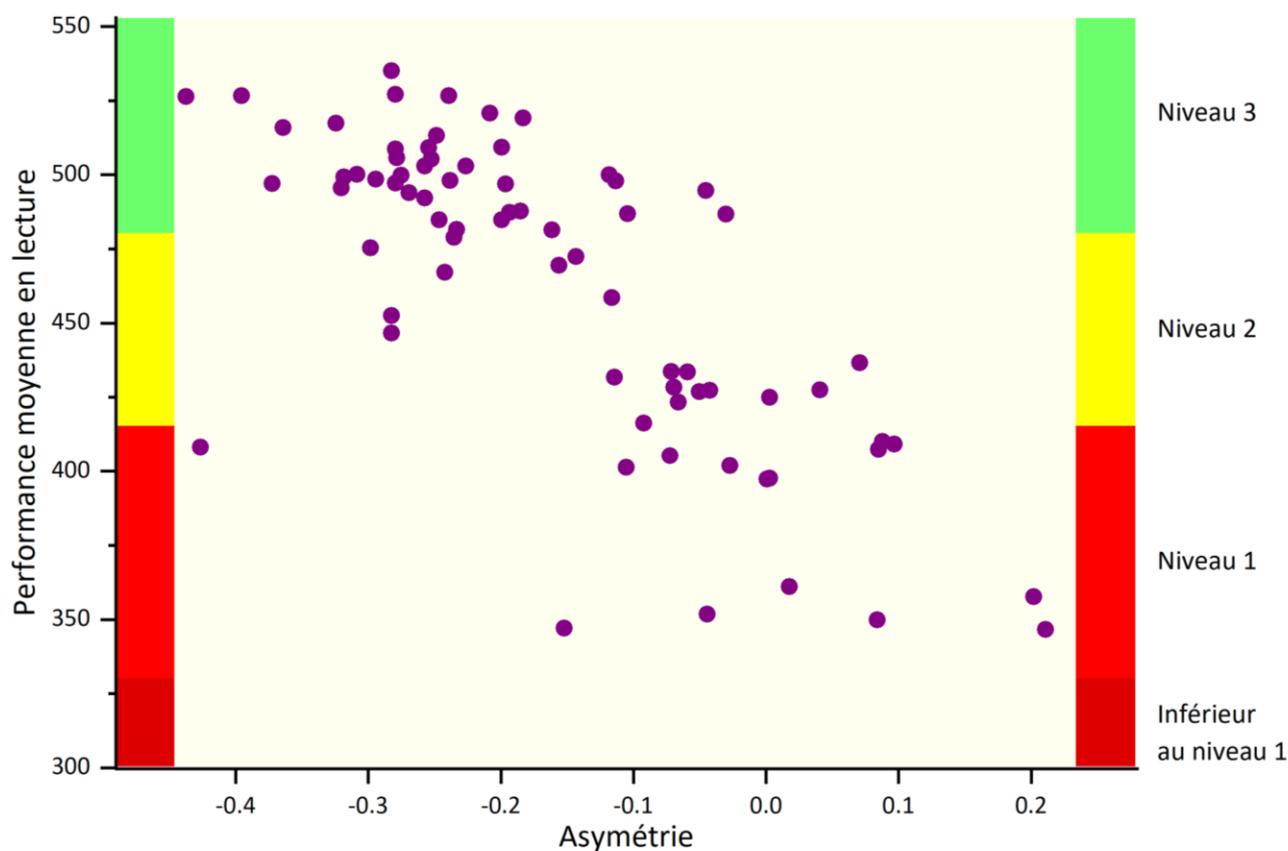
Source : PISA 2015.

Pour les mesures continues, comme les scores à des tests, l'objectif est de décrire la distribution des scores d'un petit ensemble de statistiques sommaires qui pourront être utilisées pour suivre les progrès d'une année à l'autre ou pour évaluer la performance en la comparant à une norme ou à celle d'autres territoires. Le **figure 8** montre la distribution des scores pour la République dominicaine, le Mexique et la Suède. Les deux marqueurs les plus souvent utilisés pour suivre la progression sont la moyenne et la médiane. Toutefois, trois autres statistiques sont également utiles : l'écart type, l'asymétrie et le pourcentage d'élèves « vulnérables ». L'écart type est un indicateur de la variation dans un ensemble de résultats (ou de l'étalement). Les écarts interquartiles (c'est-à-dire les valeurs comprises entre les 25^e et 75^e percentiles) des diagrammes présentés précédemment pour les scores à l'ÉPE et au NAPLAN peuvent être utilisés à cette fin. L'asymétrie indique dans quelle mesure la distribution est asymétrique. Dans les distributions qui



présentent une asymétrie négative, il y a davantage de scores sous la moyenne qu'au-dessus de la moyenne. Dans le cas inverse, il faudrait parler de distributions à asymétrie positive.

Figure 9. Relation entre la performance moyenne en lecture et l'asymétrie



Source : PISA 2015.

L'asymétrie est importante parce que, à mesure qu'ils se développent, les systèmes scolaires n'agissent pas sur toute la distribution des scores, laquelle peut présenter une asymétrie positive ou même être symétrique (c.-à-d. non asymétrique). En général, ils agissent plutôt sur les élèves qui réussissent le mieux, laissant derrière une sous-population vulnérable. Dans la figure 8, les scores de la République dominicaine ont une asymétrie positive (asymétrie = 0,20), ceux du Mexique sont presque symétriques (asymétrie = -0,07) et ceux de la Suède ont une asymétrie négative (asymétrie = -0,31). La **figure 9** montre la relation entre les scores moyens en lecture et l'asymétrie de la distribution pour les pays qui ont participé au PISA 2015. La corrélation entre les scores moyens en lecture et l'écart type est très faible (0,12), mais la corrélation entre les scores moyens en lecture et l'asymétrie est de -0,74. La Jordanie est une valeur aberrante dans cette analyse avec un score moyen faible (408) et une asymétrie négative (-0,43). Si l'on exclut cette valeur, la corrélation entre les scores moyens et l'asymétrie est de -0,80. La majorité des pays ayant des scores faibles sont des pays à revenu faible et moyen (voir la figure 5), lesquels ont en général de faibles taux de transition vers le premier



cycle du secondaire. S'il existait des résultats en lecture pour les enfants non scolarisés, la relation négative entre les scores moyens et l'asymétrie serait sans aucun doute beaucoup plus forte.

Se concentrer sur la vulnérabilité. Quand il est question de faire le suivi du développement des enfants, le débat fait rage : faut-il mettre l'accent sur les forces des enfants ou sur leurs difficultés? (Ben-Arieh et Frønes, 2007). Les études portant sur le développement des enfants s'intéressent surtout à leurs difficultés (p. ex. problèmes de santé ou de comportement). Les efforts pour évaluer l'efficacité des écoles et des systèmes scolaires ont surtout porté sur les forces des élèves en mesurant leur réussite. Certains efforts ont cependant mesuré les difficultés en lecture et le décrochage scolaire.

Parallèlement au débat sur les forces et les difficultés, les chercheurs s'interrogent également à savoir s'il est préférable d'utiliser des mesures dichotomiques ou des mesures continues. Pour les politiques, on préfère souvent les mesures dichotomiques parce qu'elles attirent l'attention sur les enjeux socioéconomiques qui sont pertinents pour les enfants vulnérables. Ces mesures sont également plus faciles à comprendre. Par exemple, un énoncé du type « la prévalence des jeunes souffrant de dépression a augmenté de 8 % à 10 % » est généralement plus facile à interpréter qu'un énoncé du type « le score moyen de la dépression a augmenté de 8,5 à 8,9 sur une échelle de 10 points ».

Le cadre de prospérité éducative privilégie les mesures des forces parce que son objectif général est de donner à tous les enfants la chance de s'épanouir. Pour les décideurs, il peut être plus facile de rallier le public autour d'un objectif comme celui d'augmenter le nombre d'enfants d'âge scolaire en bonne santé physique qu'autour d'un objectif comme celui de réduire l'obésité chez les enfants. Les objectifs du Programme de développement durable à l'horizon 2030 (Institut de statistique de l'UNESCO, 2017) sont rédigés de manière positive. Il est toutefois préférable de parler de difficultés lorsqu'il faut aborder des marqueurs comme les problèmes de santé mentale ou physique. Souffrir d'anxiété et de dépression n'est pas la même chose que d'avoir un faible résultat par rapport à un indice du bonheur. Il peut également être plus facile pour les décideurs de se concentrer sur les élèves qui ont de faibles résultats en lecture que sur ceux qui ont atteint une norme.

Par ailleurs, les mesures continues sont en général préférées aux mesures dichotomiques parce qu'elles contiennent plus d'information. Un pays peut toutefois réussir à réduire le nombre d'enfants avec de faibles scores en lecture sans que cela paraisse dans la moyenne des scores. De plus, dans de nombreux pays qui ont participé à des évaluations internationales, un nombre important d'élèves ont obtenu des scores proches ou égaux au seuil des tests. Il est même possible que les « scores réels » de ces élèves aient été inférieurs à leurs scores estimés (Nonoyama-Tarumi et Willms, 2010). Dans de tels cas, il est préférable de se concentrer sur la vulnérabilité. Certaines mesures ont des points de découpage bien établis pour définir la vulnérabilité, comme le faible poids à la naissance ou l'obésité infantile. Les études internationales sur l'éducation comprennent des points de découpage pour les valeurs de référence ou les niveaux de scolarité. Pour déterminer si un point de découpage est approprié, il faut tenir compte des résultats en cause et de leur distribution dans la population visée.



Selon les données du PISA 2015, le score moyen en lecture du Mexique est de 423. L'écart type est de 78 et l'asymétrie de -0,07. Le pourcentage d'élèves considérés vulnérables, c'est-à-dire ceux qui n'ont pas atteint le niveau 3, est de 76,0 %. De manière générale, les grandes études internationales échantillonnent d'abord des écoles, puis des élèves de chacune de ces écoles. Par ailleurs, les échantillonnages de la majorité des pays sont stratifiés pour s'assurer que certaines régions ou certains secteurs sont adéquatement représentés. Ces études font passer des tests en rotation lors desquels les élèves sélectionnés ne réalisent qu'une partie des tâches. Par conséquent, pour estimer les statistiques, il faut utiliser un ensemble de *poids de rééchantillonnage* afin de tenir compte du plan d'échantillonnage. Quand les statistiques portent sur les scores aux tests, il faut également utiliser des *valeurs plausibles* pour tenir compte du plan d'échantillonnage. Les techniques pour l'estimation de ces statistiques et de brèves descriptions des poids de rééchantillonnage et des valeurs plausibles sont fournies à l'annexe 2.

Qui est vulnérable?

Gradients socioéconomiques. Un *gradient socioéconomique*, ou « niveau d'apprentissage », décrit le lien entre un résultat social et le SSÉ de personnes d'un endroit donné (p. ex. une école, une communauté, une province, un état ou un pays) (Willms, 2003a; 2006). Le résultat social peut porter sur différents aspects, comme l'état de santé des enfants, des scores à des tests ou le niveau de scolarité. Le SSÉ désigne généralement la position qu'occupe une famille ou une personne par rapport à une structure hiérarchique établie en fonction de la richesse, du prestige et du pouvoir (Mueller et Parcel, 1981). Les mesures du SSÉ sont importantes pour le suivi de l'éducation parce qu'elles fournissent un contexte en fonction duquel établir des objectifs et mesurer les progrès vers ceux-ci. Ces mesures aident à définir une approche pour évaluer l'égalité des résultats et l'équité de l'offre scolaire, pour imaginer des stratégies potentielles dans le but d'améliorer les résultats des élèves et de réduire les inégalités, et pour renforcer la validité des études de recherche (Willms et Tramonte, 2018).

La **figure 10** montre le gradient socioéconomique des scores en lecture des élèves de 15 ans au Mexique (selon les données du PISA 2015). La mesure du SSÉ est la mesure composite utilisée par le PISA, soit l'Indice de statut économique, social et culturel. Un gradient socioéconomique comporte trois composantes : le niveau, la pente et la force de la relation entre le résultat et le SSÉ.

- a. Le *niveau* du gradient est défini comme étant le score attendu pour une personne de SSÉ moyen. Le niveau d'un gradient pour un pays (ou une province, un État ou une école) est un indicateur de sa performance moyenne après avoir pris en compte le statut socioéconomique des élèves. Le niveau du gradient du Mexique est de 450,3. Pour la présente analyse, la mesure du SSÉ a été échelonnée afin que zéro représente la moyenne de l'OCDE. Par conséquent, le niveau fait référence au score attendu pour un élève dont le SSÉ est comparable à celui d'un élève hypothétique dont le SSÉ serait égal à la moyenne de l'OCDE.
- b. La *pente* du gradient indique dans quelle mesure l'inégalité est attribuable au SSÉ. Plus la pente d'un gradient est inclinée, plus le SSÉ a une incidence sur la performance (et plus les inégalités sont grandes). Inversement, moins la pente d'un gradient est inclinée, moins le SSÉ a une incidence sur la



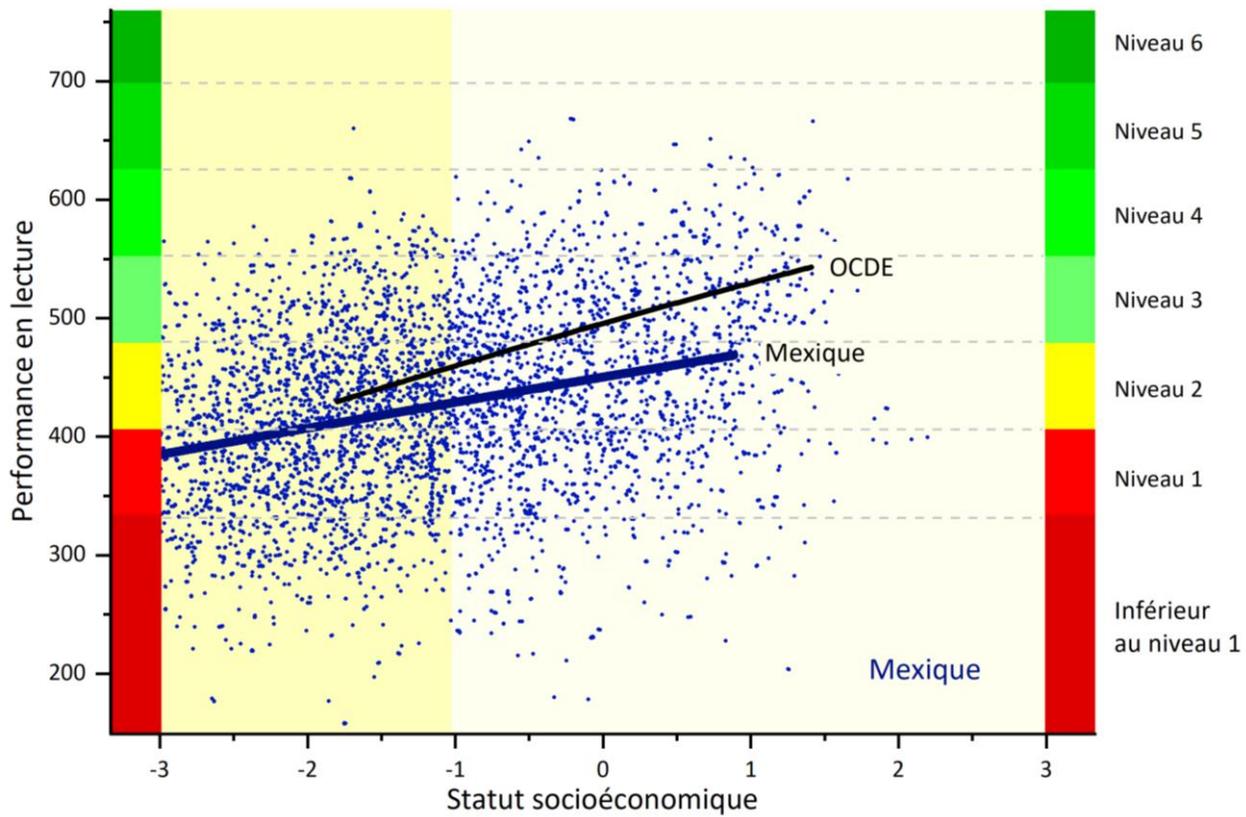
performance (et moins les inégalités sont grandes). La pente du gradient du Mexique est de 21,2 (dans le centre des données), ce qui indique que la performance attendue en lecture augmente de 21,2 points pour chaque augmentation du SSÉ équivalent à l'écart type. Le gradient du Mexique est légèrement curviligne, avec une pente qui diminue quelque peu à mesure qu'augmente le SSÉ. Toutefois, le coefficient de la composante curviligne n'est pas statistiquement significatif.

- c. La *force* du gradient est définie comme étant la proportion de la variance observée dans le résultat social qui est expliquée par le SSÉ. Si la force de la relation est grande, c'est qu'une part considérable de la variation observée dans la mesure du résultat est associée au SSÉ. En revanche, si la relation est faible, c'est qu'une part relativement petite de la variation est associée au SSÉ. La mesure la plus courante de la force d'une relation est le R au carré. Celle-ci est de 0,12 pour le Mexique.

La ligne bleue représente le gradient du Mexique. Celle-ci est tracée du 5^e au 95^e percentile pour le SSÉ. Pour le Mexique, les 5^e et 95^e percentiles sont respectivement -3,08 et 0,88. Ce sont donc 90 % des élèves mexicains qui se situent dans cet intervalle. Le SSÉ moyen des élèves du Mexique est inférieur à celui des pays de l'OCDE. Le graphique montre également la performance en lecture et le SSÉ pour un échantillon représentatif de 5 000 élèves mexicains. Ces valeurs sont représentées par les points bleus placés de part et d'autre du gradient. Ces points montrent qu'il y a une variation considérable de la performance en lecture, quel que soit le SSÉ. La ligne noire représente le gradient pour tous les pays de l'OCDE. Pour les élèves de l'OCDE, le 5^e et le 95^e percentiles sont respectivement de -1,78 et de 1,40.

Le terme « niveau d'apprentissage » est une métaphore du gradient socioéconomique. La question à laquelle toutes les écoles sont confrontées est la suivante : « Comment pouvons-nous augmenter le niveau d'apprentissage et diminuer son inclinaison ? ». Plusieurs approches peuvent être adoptées afin d'augmenter la performance et de réduire les inégalités entre les élèves de différents SSÉ. La meilleure approche à adopter dépendra des enjeux sociaux et politiques, mais également de la distribution des résultats des élèves et du SSÉ au sein des écoles et d'une école à l'autre, et de la relation et de l'interaction de ces facteurs avec les bases de la réussite.

Figure 10. Gradient socioéconomique de la performance en lecture pour le Mexique



Source : PISA 2015.



Rendement décroissant. Selon l'hypothèse du « rendement décroissant », les effets du SSÉ sur les résultats sociaux sont plus faibles pour les SSÉ plus élevés. Par exemple, on pourrait avancer que, à partir d'un certain niveau, le SSÉ n'a que peu ou pas d'incidence sur la performance en lecture. Cela ne semble toutefois pas le cas au Mexique puisque la ligne du gradient pour la performance en lecture augmente de façon linéaire à mesure que le SSÉ augmente. L'estimation de la ligne du gradient comprend un terme quadratique pour le SSÉ. Il s'agit de la statistique utilisée pour évaluer si le rendement associé au SSÉ augmente ou diminue pour les niveaux les plus élevés de SSÉ. Pour le Mexique, le coefficient pour le terme linéaire (c.-à-d. la pente) est de 21,2, alors que le terme quadratique est de -0,3. Il y a donc un rendement décroissant pour les niveaux les plus élevés de SSÉ, mais celui-ci n'est pas statistiquement significatif.

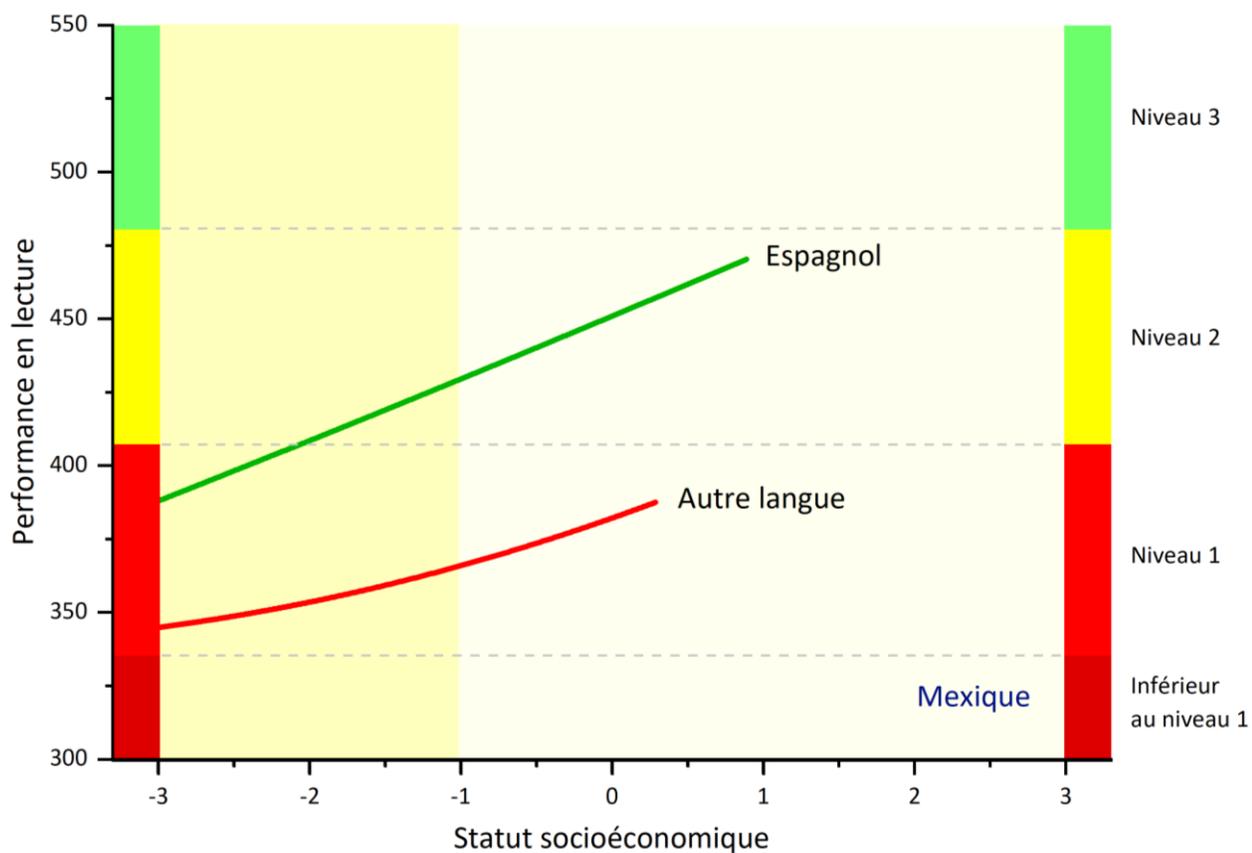
Les techniques employées pour l'estimation des statistiques qui ont permis d'établir les gradients socioéconomiques sont décrites à l'annexe 2.

Willms et Somers (2001) ont observé des rendements *croissants* associés au SSÉ en lecture et en mathématiques chez les élèves de 3^e et de 4^e année de plusieurs pays d'Amérique latine. Ces auteurs ont suggéré que les enfants dont les parents avaient terminé l'école secondaire étaient avantagés. Pour les élèves de 15 ans, un rendement croissant peut indiquer que ceux-ci n'ont pas réussi leur transition de « apprendre à lire » à « lire pour apprendre » à la fin du premier cycle du primaire et qu'ils n'ont pas pu profiter autant de l'enseignement au cours des années qui ont suivi. Un rendement croissant peut également être attribuable à un effet de plancher du test du PISA. En effet, dans la majorité des pays à revenu faible et moyen, un important pourcentage d'élèves de faible SSÉ ne parviennent pas à répondre même aux questions les plus faciles du test.

Égalité. La différence entre les scores de deux sous-populations à des tests est une mesure directe de l'égalité. La différence entre les scores en lecture des élèves dont la langue maternelle est l'espagnol et ceux des élèves dont la langue maternelle n'est pas l'espagnol est de 74 points environ. L'approche du gradient socioéconomique permet d'évaluer l'ampleur de la différence après avoir pris en compte le SSÉ. Elle permet de déterminer si le SSÉ a un impact significatif (c.-à-d. si l'écart change selon le niveau de SSÉ). La **figure 11** montre les lignes des gradients pour ces deux sous-populations. Dans ce cas, le SSÉ n'explique que partiellement la différence entre les deux sous-populations. Pour les élèves de SSÉ moyen, cette différence est de 69 points environ. La figure montre également que la différence entre les deux sous-populations est presque toujours la même, quel que soit le SSÉ.



Figure 11. Gradients socioéconomiques de la performance en lecture selon la langue maternelle au Mexique



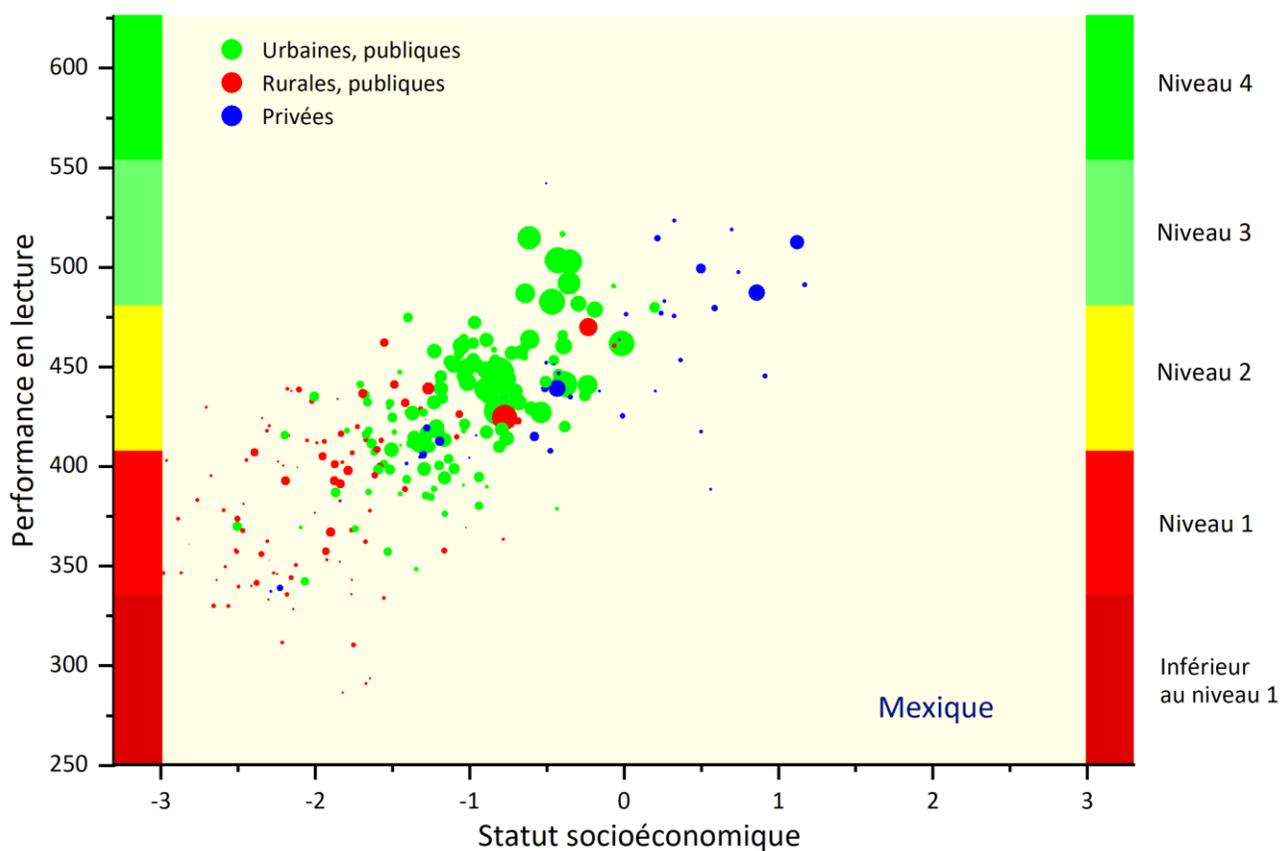
Source : PISA 2015.



Où sont les enfants vulnérables?

Profil des écoles. Un profil des écoles est tout simplement un diagramme de distribution des scores moyens des écoles par rapport à leur SSÉ moyen pour un résultat. La **figure 12** montre le profil des écoles pour les scores en lecture au Mexique. Chaque point représente une des écoles qui ont participé au PISA 2015. La taille des points est proportionnelle à la racine carrée de l'effectif d'une école. Les écoles publiques rurales sont représentées par des points rouges, les écoles publiques urbaines par des points verts et les écoles privées par des points bleus.

Figure 12. Profil des écoles pour la performance en lecture au Mexique



Source : PISA 2015.

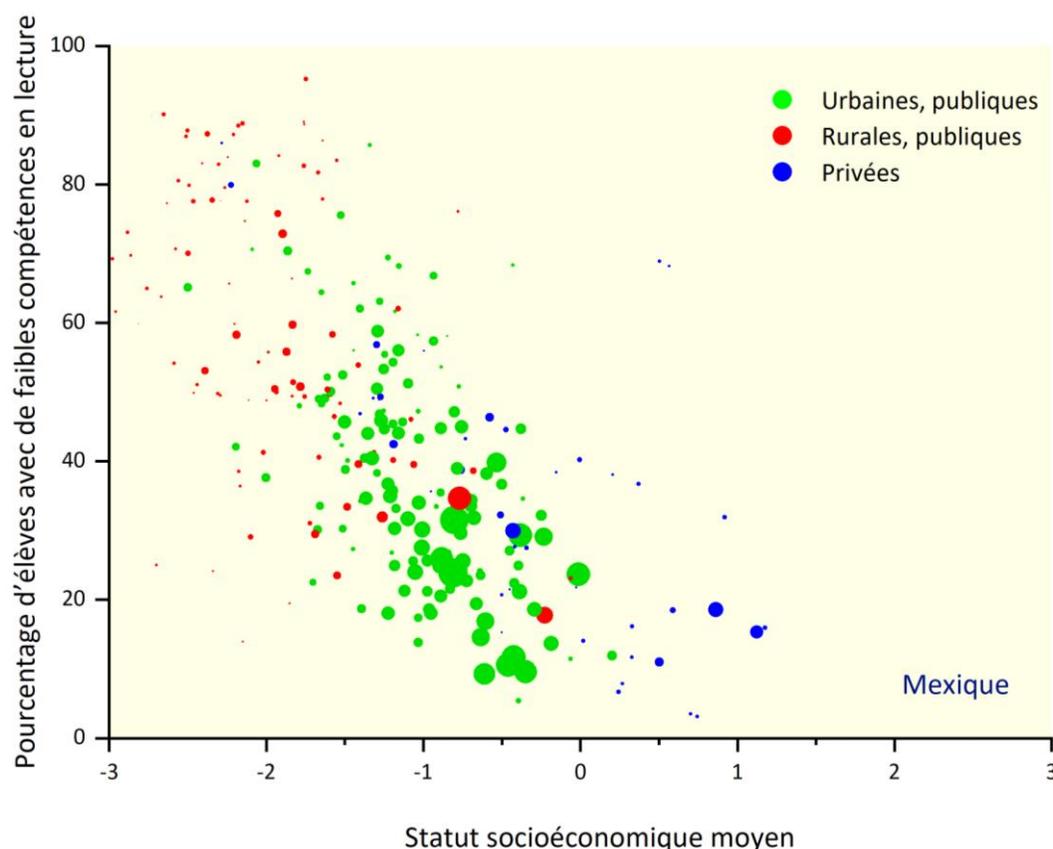
Trois choses sautent immédiatement aux yeux lorsque l'on regarde le profil des écoles.

- a. Les scores moyens en lecture des écoles sont très différents. À tous les niveaux de SSÉ, l'écart entre les écoles avec les scores les plus faibles et celles avec les scores les plus élevés est d'environ 100 points. Ceci correspond à un écart type ou à une ampleur de l'effet de 1,00.



- b. Si on établit le « seuil de la pauvreté » à un SSÉ de -1,0, alors presque les deux tiers (65 %) des écoles ont un SSÉ moyen inférieur au seuil de la pauvreté. La majorité de ces écoles sont petites et situées en milieu rural. Une mesure plus détaillée du SSÉ de même qu'une mesure de la pauvreté ont été élaborées pour le PISA pour le développement (Tramonte et Willms, 2018).
- c. Le SSÉ moyen de la majorité des écoles privées est beaucoup plus élevé que celui des écoles publiques urbaines. Toutefois, les scores moyens en lecture de ces écoles privées sont comparables à ceux des écoles publiques urbaines dont le SSÉ moyen est supérieur à -0,5.

Figure 13. Profil des écoles pour les élèves ayant de faibles compétences en lecture (niveau 1 et inférieur) au Mexique



Source : PISA 2015.

La **figure 13** présente le profil des écoles pour les élèves ayant de faibles compétences en lecture au Mexique. Pour cette analyse, il a été établi que les élèves qui avaient obtenu le niveau 1 ou moins au PISA avaient de faibles compétences en lecture. Ce type de profil, qui montre le « pourcentage de vulnérabilité », est utile pour les résultats dichotomiques. Ces résultats peuvent être négatifs (p. ex. faible niveau de scolarité ou problèmes de santé mentale) ou positifs (p. ex. achèvement scolaire). Ce profil va de pair avec le graphique de concentration de la vulnérabilité présenté plus loin.



Systèmes scolaires inclusifs. Les écoles des systèmes scolaires inclusifs permettent à tous les enfants de s'épanouir. Tous les enfants, peu importe leur genre, leur classe sociale, leur ethnie, leur nationalité, leur orientation sexuelle, leur religion ou s'ils ont un handicap (Willms, 2009a). Pour que les enfants puissent s'épanouir, on doit pouvoir assurer leur bien-être scolaire, social, émotionnel, physique et spirituel. L'inclusion fait référence à un ensemble de valeurs et de croyances selon lesquelles tous les enfants ont le droit en vertu de la loi de recevoir une éducation dans un environnement sûr aux côtés de leurs voisins du même âge. L'inclusion suppose « une vision commune qui englobe tous les enfants de la tranche d'âge concernée, et la conviction qu'il est de la responsabilité du système éducatif général d'éduquer tous les enfants » (UNESCO, 2005, p. 13). Pour atteindre l'inclusion, il faut prendre des dispositions pour aider les élèves qui ont des besoins particuliers ou des problèmes de comportement. Toutefois, le terme inclusion signifie de manière générale que tous les enfants ont le droit de recevoir une éducation dans l'environnement le moins restrictif possible. Pour développer un système scolaire inclusif, il faut que les parents, les éducateurs, les politiciens et les autres leaders des communautés prennent position en faveur de l'inclusion en luttant contre les attitudes discriminatoires et en soutenant les initiatives visant à accroître l'intégration d'apprenants aux besoins variés (Riehl, 2000).

Dans les systèmes scolaires inclusifs, les résultats des élèves sont meilleurs et il y a moins d'inégalités (Willms, 2010). Dans les systèmes scolaires inclusifs, les bases de la réussite (enseignement de qualité, temps d'apprentissage, ressources matérielles et soutien des familles et de la communauté) sont réparties plus également entre les écoles. Dans ces systèmes, les élèves sont eux-mêmes des ressources clés. Les élèves qui ont de faibles capacités ont de meilleures chances de réussir lorsqu'ils étudient dans des classes hétérogènes aux côtés de pairs qui ont des attentes élevées et qui sont intellectuellement curieux.

Inclusion horizontale. À l'échelle d'un système scolaire, on peut parler de deux types d'inclusion : l'inclusion horizontale et l'inclusion verticale. Dans les systèmes scolaires horizontalement inclusifs, la variation entre les SSÉ est relativement faible. Au Mexique, selon le profil des écoles, l'étendue des SSÉ moyens va de -3,0 à 1,2. La ségrégation résidentielle est le principal facteur contribuant à la « ségrégation horizontale ». Au Mexique, par exemple, il y a de grandes différences socioéconomiques entre les quartiers des grandes villes et entre les régions urbaines et les régions rurales.

Inclusion verticale. Dans un système scolaire verticalement inclusif, la variation entre les résultats des élèves des écoles ayant des SSÉ similaires est faible. Le profil des écoles pour le Mexique montre que l'étendue des scores en lecture de la majorité des écoles va de 350 à 550. Dans les systèmes scolaires où il y a de la « ségrégation verticale » (c.-à-d. qui ne sont pas verticalement inclusifs), les élèves sont dirigés vers certains types d'écoles ou de programmes en fonction de leurs résultats scolaires préalables ou sur la base d'une mesure de leurs capacités cognitives ou langagières. Les élèves peuvent également être dirigés vers des groupes ou des programmes spéciaux de leur école, comme les groupes destinés aux élèves à besoins particuliers. Le redoublement, les programmes d'immersion linguistique et le regroupement des élèves selon leurs capacités sont des exemples de facteurs qui contribuent à la ségrégation verticale.



Les écoles privées contribuent à la ségrégation verticale et à la ségrégation horizontale. Si elles sélectionnent leurs élèves sur la base de leurs résultats scolaires préalables, elles contribuent à la ségrégation verticale. Si elles exigent des droits de scolarité élevés, elles deviennent inaccessibles aux familles de faible SSÉ et contribuent ainsi à la ségrégation horizontale.

Deux statistiques sont employées pour résumer le niveau d'inclusion verticale et horizontale. La variation dans la performance des élèves est présentée de deux manières : la variation intra-école et la variation inter-écoles. L'inclusion verticale correspond au degré de variation *intra-école* d'un résultat. La variation du SSÉ des élèves est également présentée de deux manières : la variation intra-école et la variation inter-écoles. L'inclusion horizontale correspond au degré de variation *intra-école* du SSÉ. En théorie, ces indices peuvent aller de 0 % à 100 %, mais ils vont de 10 % à 50 % pour la majorité des pays (Willms, 2010). Les techniques statistiques employées pour estimer ces indices sont décrites à l'annexe 2.

Prenez note que la mesure de l'inclusion verticale dépend du résultat étudié. Comme le niveau moyen de bien-être des élèves est assez similaire d'une école à l'autre, les estimations de l'inclusion verticale basées sur une mesure comme l'estime de soi produiront un indice d'inclusion verticale beaucoup plus élevé que celles basées sur une mesure comme la performance en lecture. Compte tenu de l'importance des compétences en lecture, il serait utile de produire un indice de l'inclusion verticale basé sur un test de lecture national ou international pour faire le suivi de l'inclusion verticale. Le degré d'inclusion verticale peut également varier en fonction du degré de scolarité, surtout dans les systèmes scolaires qui sélectionnent les élèves sur la base de leurs résultats scolaires ou de leurs compétences à certains âges.

Concentration de la vulnérabilité. Le graphique de la concentration de la vulnérabilité prend la forme d'une courbe de Lorenz, laquelle est utilisée en économie pour représenter la distribution des revenus ou de la richesse (Lorenz, 1905). Le graphique de concentration de la vulnérabilité montre la concentration d'élèves vulnérables dans les écoles d'un pays. La **figure 14** en est un exemple. Selon le PISA 2015, il y a au Mexique 44 335 écoles fréquentées par des jeunes de 15 ans inscrits en 7^e année ou à un niveau supérieur. Ceci représente 1 392 995 élèves. Parmi ces élèves, 41,8 % ont des compétences en lecture de niveau 1 ou inférieur. L'histogramme présenté à la figure 14 compte 443 barres, chacune représentant un ensemble de 100 écoles. Chaque barre représente le nombre estimé d'élèves vulnérables dans chaque ensemble d'écoles. Les barres sont classées en fonction du nombre d'élèves vulnérables dans l'ensemble de 100 écoles. Par exemple, la 400^e barre regroupe 3 826 élèves de 100 écoles, ce qui représente environ 38 élèves en moyenne par école avec un niveau 1 ou inférieur en lecture. La barre la plus élevée regroupe 10 838 élèves, ce qui représente environ 108 en moyenne par école.

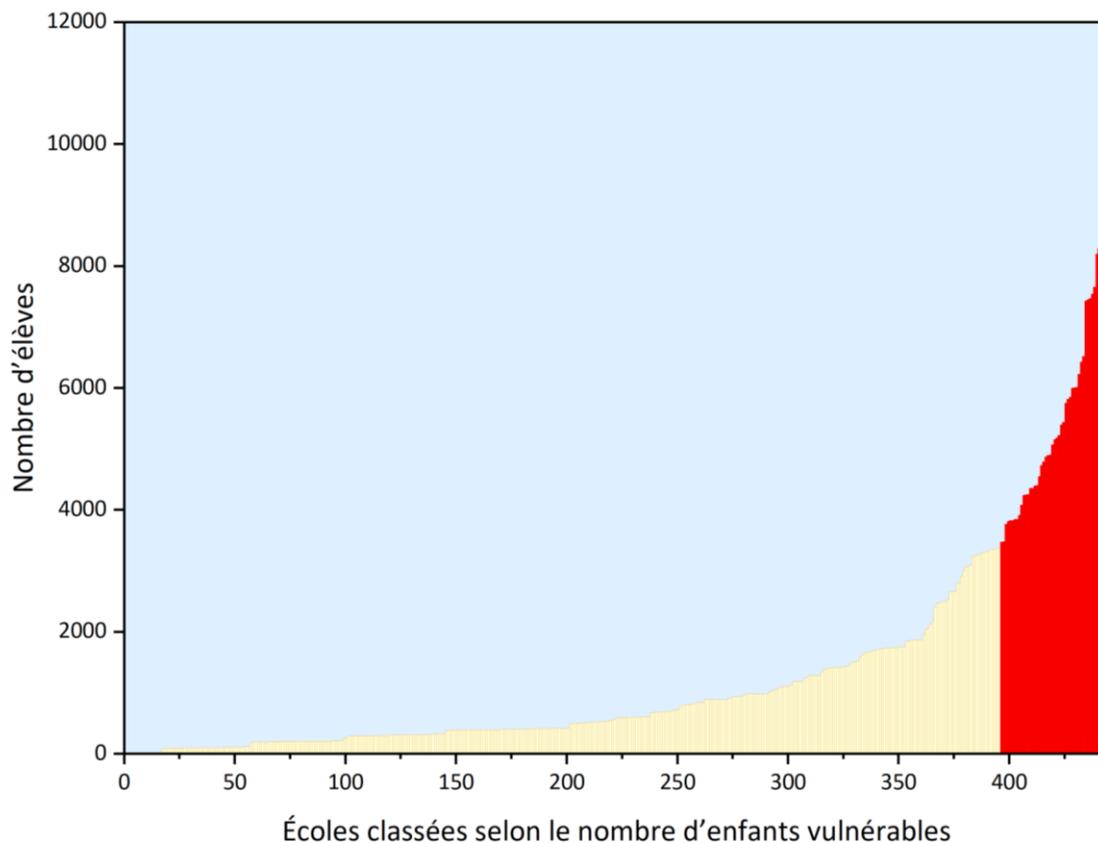
Ce graphique révèle que 50 % des élèves ayant de faibles compétences en lecture sont concentrés dans 11 % des écoles, ce qui représente moins de 5 000 écoles. Par ailleurs, 75 % des élèves vulnérables sont concentrés dans 18 % des écoles, ce qui représente un peu moins de 8 000 écoles. Parmi les cinq types de stratégies présentées à la section 4 du présent rapport, il y a celles qui ciblent les résultats. Dans le cas présent, il pourrait être pertinent de mettre en place une telle stratégie dans 5 000 à 8 000 écoles. Cette



tâche peut paraître colossale, mais il faut avoir à l'esprit qu'il y a plus de 44 000 écoles fréquentées par des élèves de 15 ans au Mexique.



Figure 14. Concentration d'élèves avec de faibles compétences en lecture dans les écoles du Mexique



Source : PISA 2015.

Un graphique de concentration de la vulnérabilité peut également permettre d'évaluer la concentration dans n'importe quelle sous-population (p. ex. les élèves autochtones, les élèves ayant un handicap). Il n'est pas nécessaire que la sous-population soit « vulnérable ». Par exemple, au Mexique, environ 4,5 % des élèves ont des compétences en lecture de niveau 4 ou supérieur – 50 % de ces élèves sont concentrés dans moins de 2 % des écoles du pays.

Gradients socioéconomiques intra-école et inter-écoles. Le gradient socioéconomique présenté à la figure 10 est le gradient global pour le Mexique. Il est possible d'estimer un gradient pour n'importe quel territoire (p. ex. un pays, une région, un district scolaire). Toutefois, chaque école, où qu'elle soit, a un gradient socioéconomique qui lui est propre et qui a un niveau, une pente et une force. Un modèle multiniveaux décrit les gradients de toutes les écoles simultanément et les relations entre elles. Au Mexique, le gradient moyen intra-école est de 7,7 et le gradient moyen inter-écoles est de 41,0. Ces données suggèrent que les élèves d'une même école ont souvent des compétences en lecture de niveau similaire. Les inégalités dans les compétences en lecture associées au SSÉ sont surtout observées entre les écoles.



Composition de l'école. L'effet de la composition de l'école est l'incidence qu'ont les caractéristiques d'une école (p. ex. SSÉ moyen, pourcentage d'élèves ayant un handicap, pourcentage d'élèves qui parlent une autre langue que la langue d'enseignement à la maison) sur les résultats des élèves (Alexander et Eckland, 1975; Bryk et Driscoll, 1988; Willms, 1986; Willms, 2010). Cet effet s'ajoute aux effets associés au SSÉ. Un effet de la composition positif pour le SSÉ moyen indique que les élèves qui fréquentent des écoles de SSÉ élevé ont généralement de meilleurs résultats que ceux qui fréquentent des écoles de faible SSÉ, même après avoir pris en compte les SSÉ individuels des élèves. L'effet de la composition associé au SSÉ moyen de l'école est de 35,2 pour le Mexique.

L'effet de la composition doit être considéré comme une approximation pour un certain nombre de facteurs qui sont corrélés à la composition de l'école parce que, avec les données transversales, il est impossible de distinguer les effets associés à la composition de l'école de ceux attribuables aux facteurs de base ou à d'autres processus scolaires (Raudenbush et Willms, 1995). Les effets de la composition peuvent en partie être attribuables aux pairs. Par exemple, lorsque des élèves intelligents et motivés travaillent ensemble, ils apprennent les uns des autres et se fixent des objectifs plus élevés (Robertson et Symons, 1996; Zimmer et Toma, 1997). Par ailleurs, de manière générale, les écoles de SSÉ élevé attirent et retiennent des enseignants talentueux et motivés, ont davantage de ressources pédagogiques et reçoivent un meilleur appui des parents (Willms, 1986; Willms et Somers, 2001; OCDE, 2001). On pourrait dire également que les bases de la réussite sont plus solides dans ces écoles. Le terme « effet » a une connotation causale. Il ne faut toutefois pas inférer une relation causale à partir des données d'études transversales, comme celle du PISA (Alexander, Fennessey, McDill et D'Amico, 1979).

Variabilité de l'effectif. Un certain nombre de politiques et de pratiques en éducation reposent sur un argument : lorsque les groupes et les écoles sont plus homogènes, les élèves apprennent plus rapidement. Cet argument sous-tend les politiques qui visent à orienter les élèves vers des programmes ou des écoles à vocation particulière et la pratique qui consiste à faire redoubler les élèves dont les résultats scolaires sont considérablement inférieurs à la norme pour leur niveau.

Tout comme l'effet de la composition de l'école, l'effet de la composition de l'effectif est estimé en incluant une mesure de la variabilité de l'effectif, comme l'écart type du SSÉ de l'école, comme variable au niveau de l'école dans un modèle multiniveaux. Au Mexique, la variabilité de l'effet de l'effectif est de 5,9. Cet effet est trop faible pour être statistiquement significatif.

Gradients convergents. Des recherches antérieures basées sur les données du PIRLS 2001 et du PISA 2000 ont révélé que les compétences en lecture des élèves de SSÉ élevé n'étaient pas très différentes d'un pays à l'autre comparativement à celles des élèves de faible SSÉ (Willms, 2006). En d'autres termes, les gradients tendent à converger pour les SSÉ plus élevés. Ce schéma a également été observé chez les jeunes de 16 à 25 ans qui ont participé à l'Enquête internationale sur l'alphabétisation des adultes de 1997 (Willms, 2003b). Chaque pays devrait se demander si les gradients de leurs écoles convergent eux aussi pour les SSÉ plus élevés. Si c'est le cas, c'est peut-être que les élèves de SSÉ élevé ont tendance à réussir dans la majorité des écoles, tandis que ceux de faible SSÉ réussissent de manière très variable, selon l'école qu'ils fréquentent.



La statistique pour estimer l'ampleur de la convergence est la corrélation entre les niveaux des gradients et de leurs pentes. Cette statistique peut être estimée à l'aide d'un modèle de régression à deux niveaux qui comprend une mesure du SSÉ de la famille des élèves et permet une variation inter-écoles des points de rencontre et des pentes du SSÉ. Pour le Mexique, la corrélation est de 0,37, ce qui indique que les pentes divergent à mesure que le SSÉ augmente.

L'estimation de cette statistique basée sur un modèle de gradient multiniveaux est décrite à l'annexe 2.



IV. Politiques sur les stratégies et leur mise en œuvre

Les données de suivi peuvent orienter les politiques en répondant aux questions sur la performance du système scolaire : Où en sommes-nous? Qui sont les enfants vulnérables? Où sont les enfants vulnérables? Plusieurs statistiques et graphiques pouvant être utilisés pour brosser un portrait d'un système scolaire ont été présentés dans la section précédente. Ceux-ci sont des approches fiables et cohérentes pour évaluer les niveaux de performance des élèves et l'égalité de la performance de sous-populations d'élèves. Des données de suivi peuvent également être utilisées pour établir des objectifs réalistes et mesurables afin d'améliorer les résultats des élèves et de réduire les inégalités.

Une politique éducative doit établir des objectifs et élaborer un plan d'action pour les atteindre. Le « plan d'action » doit définir un petit ensemble de stratégies qui permettront d'atteindre les résultats et établir un plan pour leur mise en œuvre. Il faut établir des priorités, fixer des cibles à court et à long terme qui sont conformes aux objectifs et suivre les progrès vers leur atteinte. Il faut également élaborer des politiques sur les meilleures façons d'affecter les ressources disponibles. Le suivi des données est primordial pour l'établissement d'un ensemble de stratégies et leur mise en œuvre.

La présente section présente trois questions à se poser sur les stratégies et sur leur mise en œuvre. Première question : « Quels types de stratégies sont le plus susceptibles de conduire à une amélioration du système scolaire? ». Deuxième question : « Comment faut-il affecter les ressources si l'on veut renforcer les bases de la réussite et bâtir un système scolaire plus équitable? ». Troisième question : « Comment les données de suivi peuvent-elles être utilisées pour évaluer les effets de politiques sur une ou plusieurs caractéristiques structurelles clés du système scolaire? ».

Cinq types de stratégies. Willms (2006) décrit cinq types de stratégies qui peuvent être mises en œuvre par un pays, une province, un état, un territoire ou une école.³ Les relations entre les résultats et le SSÉ des élèves (représentées par les gradients socioéconomiques et les profils des écoles) peuvent aider les décideurs à déterminer quelle stratégie ou combinaison de stratégies pourra mieux relever et stabiliser le niveau d'apprentissage. Il existe cinq types de stratégies. Celles-ci peuvent être universelles, cibler la performance, cibler le SSÉ, être compensatoires ou être de redistribution. Ces stratégies sont expliquées plus en détail ci-dessous avec à titre d'exemple les données du Mexique au PISA 2015. Pour chaque type de stratégies, l'effet hypothétique d'une stratégie est évalué. Dans les graphiques présentant les effets hypothétiques, la ligne rouge représente la situation avant la mise en œuvre d'une stratégie. Il s'agit donc du gradient calculé à partir des données du PISA 2015. La ligne verte représente quant à elle le gradient après la mise en œuvre de la stratégie.

Stratégies universelles. Une stratégie universelle a pour objectif d'améliorer les résultats de tous les élèves d'un territoire donné. Les stratégies suivantes sont considérées comme universelles parce qu'elles visent

³ La version antérieure de *Les écarts en matière d'apprentissage* (Willms, 2006) utilisait le terme « intervention », mais ce terme est souvent employé pour faire référence à des stratégies spécifiques, comme l'adoption d'un nouveau programme de lecture ou d'une nouvelle méthode pédagogique.



tous les élèves, peu importe leur SSÉ : refonte des programmes d'études, réduction de la taille des groupes, modification de l'âge d'entrée à la maternelle, augmentation du temps consacré à l'apprentissage de la lecture.

La **figure 15** montre l'effet d'une stratégie universelle ayant un effet uniforme sur tous les élèves, peu importe leur SSÉ. L'ampleur de l'effet de la stratégie hypothétique est de 0,5, ce qui correspond à 50 points sur l'échelle du PISA. Certaines stratégies universelles produisent un plus grand effet chez les élèves de faible SSÉ que chez ceux de SSÉ élevé (ou vice-versa).

Stratégies ciblant la performance. Une stratégie visant les élèves qui ont eu des résultats faibles est une stratégie ciblant la performance. Par exemple, l'Évaluation de la petite enfance est utilisée dans plusieurs pays pour évaluer le développement des enfants de 3 à 6 ans, alors qu'ils se préparent à faire la transition vers l'éducation formelle (The Learning Bar, 2011). Les données recueillies permettent de classer les élèves dans trois groupes, sur la base de leurs scores dans cinq domaines. Ce classement fournit de l'information aux enseignants sur le *type* et la *quantité* de soutien dont chaque élève a besoin.

Il est également possible d'instaurer une stratégie ciblant la performance à l'échelle de l'école. Par exemple, un programme de lecture peut être administré dans un échantillon d'écoles dont la performance moyenne est faible. Dans les systèmes scolaires dont l'indice d'inclusion verticale est faible, il est plus efficace de mettre en place une stratégie à l'échelle de l'école dans un petit nombre d'écoles. Un graphique de la concentration de la vulnérabilité peut être utilisé pour estimer le nombre d'enfants qui seraient atteints par une intervention menée dans un nombre donné d'écoles.

La **figure 16** montre l'effet d'une stratégie ciblant la performance mise en place pour tous les élèves de niveau 2 ou inférieur. Cette stratégie hypothétique a une ampleur de l'effet de 0,50. Dans ce cas, la stratégie a relevé et stabilisé le niveau d'apprentissage parce qu'il y a proportionnellement moins d'élèves de faible SSÉ qui ont de faibles compétences.

Stratégies ciblant les risques. Une stratégie ciblant les risques a pour objectif d'offrir du soutien ou des ressources additionnels aux enfants jugés à risque d'échec scolaire, comme les enfants de faible SSÉ. Il y a une différence entre une stratégie ciblant les risques et une stratégie ciblant les résultats. La stratégie ciblant les risques intervient auprès des enfants considérés à risque, tandis que la stratégie ciblant les résultats intervient auprès des enfants qui ont déjà de faibles résultats. Un programme préscolaire destiné aux enfants de familles à faible revenu est un bon exemple d'une stratégie ciblant les risques. Ce type de stratégie peut également cibler une sous-population considérée vulnérable, comme les nouveaux arrivants ou les élèves membres d'une minorité ethnique.

La **figure 17** est un exemple d'une stratégie ciblant les risques avec une ampleur de l'effet de 0,50 pour les élèves dont le SSÉ est inférieur à -2,0. Cette stratégie relève et maintient le niveau d'apprentissage des élèves de faible SSÉ, ce qui réduit l'écart entre les élèves de faible SSÉ et ceux de SSÉ élevé. Les stratégies ciblant les risques ne ciblent toutefois pas les élèves de SSÉ élevé qui ont de faibles résultats. La mise en œuvre d'une stratégie ciblant la performance conjointement avec une stratégie ciblant les risques permettrait de remédier à ce problème. Il n'est toutefois pas toujours possible de mettre en place une stratégie ciblant la



performance. Par exemple, pendant les premiers stades de vie, il est difficile de mesurer le développement des enfants. Il est donc plus simple de mettre de l'avant une stratégie ciblant les enfants jugés à risque. Plusieurs facteurs de risque peuvent être retenus, dont le faible SSÉ des familles, et les enfants nés prématurément ou de faible poids à la naissance. Par ailleurs, il est impossible d'établir de stratégies ciblant la performance pour certains facteurs, comme le décrochage scolaire ou les comportements sexuels à risque. Dans ces cas, il faut utiliser l'information disponible sur les facteurs de risque et de protection pour identifier la population ciblée.

Figure 15. Effets d'une stratégie universelle avec une ampleur de l'effet de 0,50

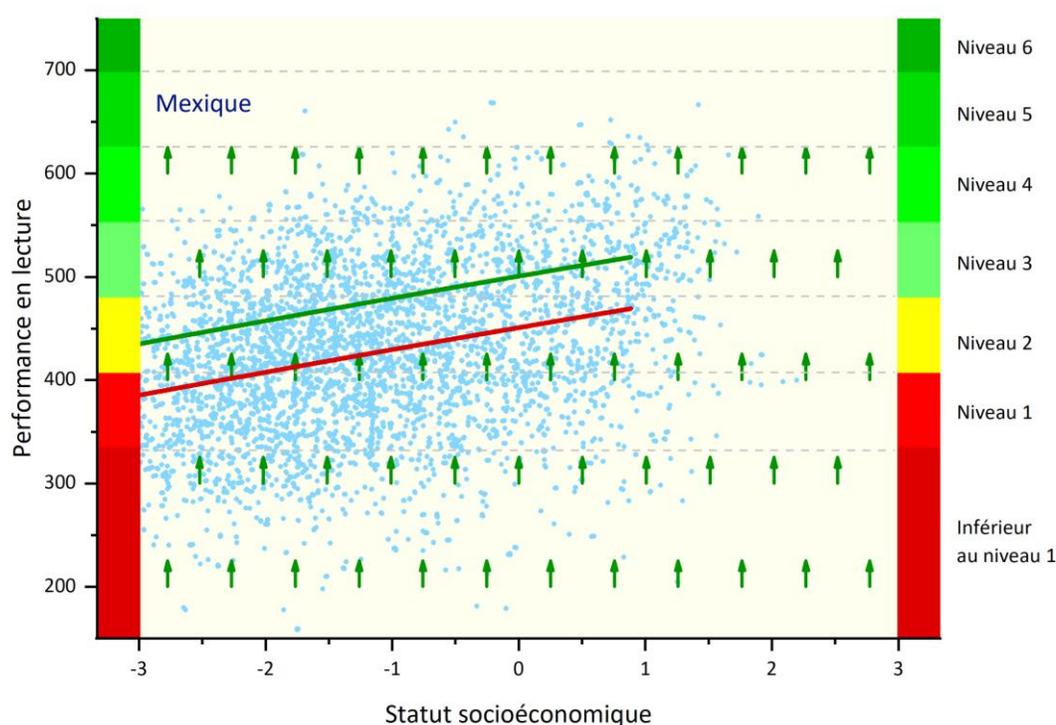


Figure 16. Effets d'une stratégie ciblant la performance avec une ampleur de l'effet de 0,50 pour les élèves ayant des compétences en lecture de niveau 1 ou inférieur

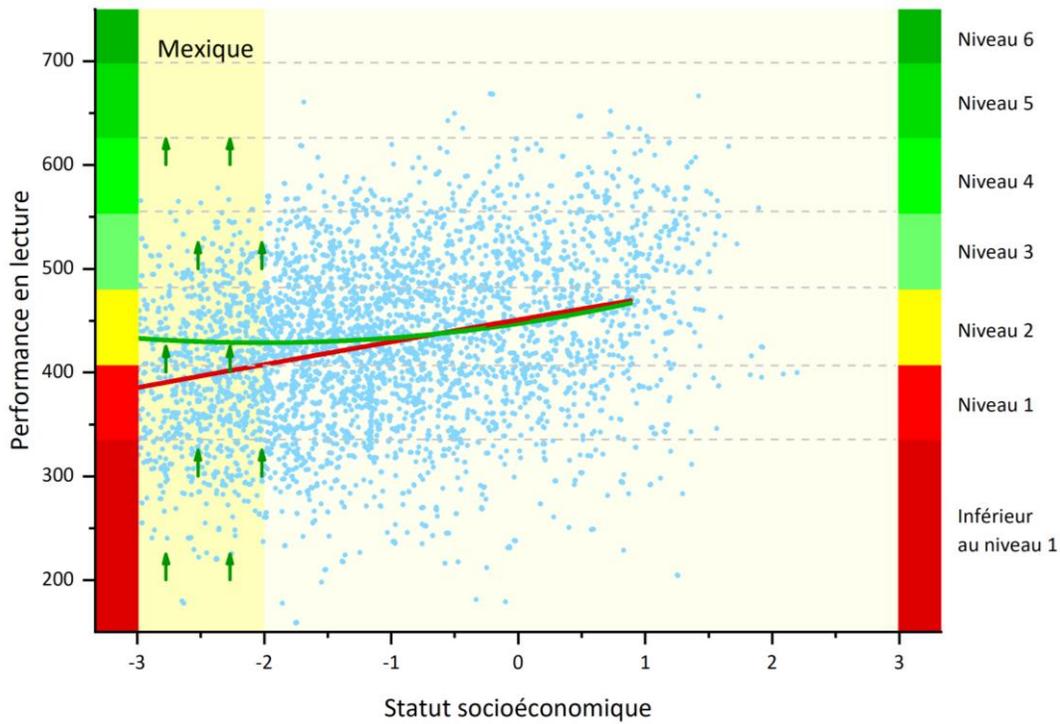
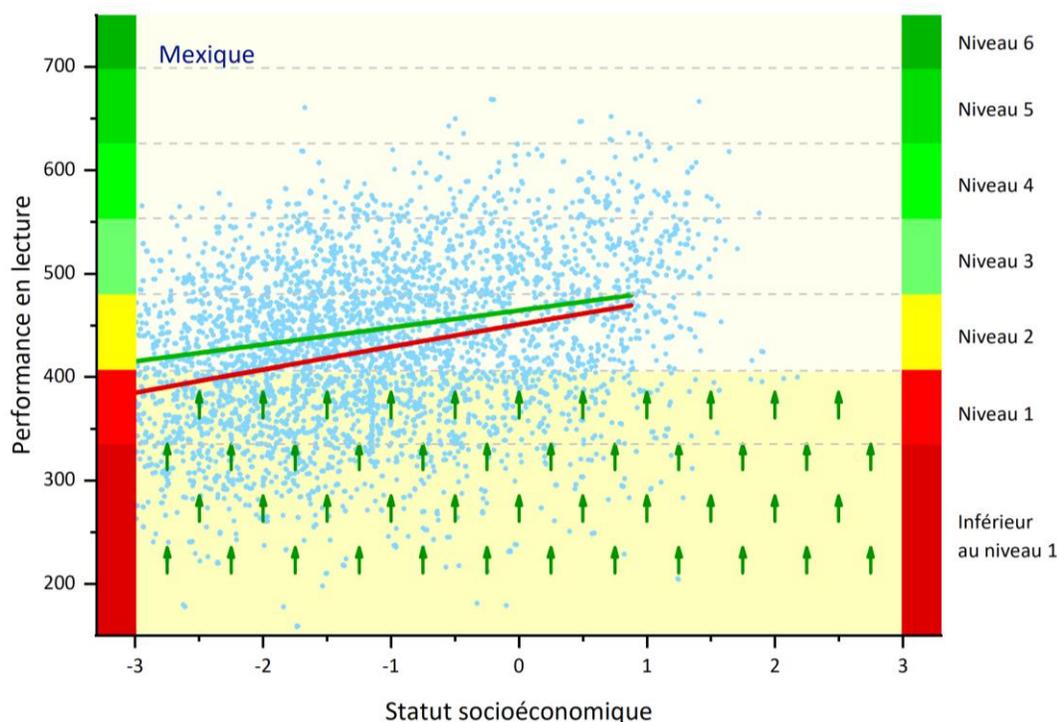


Figure 17. Effets d'une stratégie ciblant les risques avec une ampleur de l'effet de 0,50 pour les élèves dont le SSÉ est de -2 ou moins



De plus, les élèves qui sont considérés comme à risque, mais qui ont de bons résultats sont également ciblés par la stratégie. On pourrait soutenir qu'une telle stratégie n'est pas rentable. Toutefois, si ces enfants pouvaient améliorer encore davantage leurs compétences, ils pourraient peut-être devenir des modèles pour d'autres enfants au bagage familial similaire.

Stratégies compensatoires. Une stratégie compensatoire fournit des ressources pédagogiques additionnelles aux élèves de faible SSÉ ou aux élèves jugés « à risque » pour d'autres raisons. Le terme « à risque » peut être employé pour faire référence aux enfants qui n'ont pas atteint un résultat de développement ou qui risquent de manière plus générale de ne pas atteindre plusieurs résultats. La sous-population visée peut être la même que celle visée par une stratégie ciblant le SSÉ. Toutefois, l'objectif d'une stratégie compensatoire est d'agir positivement sur le SSÉ des enfants de manière à améliorer leurs résultats scolaires. Les programmes qui donnent aux élèves de faible SSÉ des manuels scolaires ou le repas du matin ou du midi sont des exemples de stratégies compensatoires.

La **figure 18** montre les effets d'une stratégie compensatoire qui augmente le niveau de SSÉ de 0,25 d'un écart type. Cette stratégie hypothétique permettrait d'améliorer les compétences en lecture des enfants ciblés d'une manière proportionnelle à la pente du gradient socioéconomique. Comme la pente du gradient du Mexique est de 21,2, cette stratégie permettrait d'augmenter les scores de 21,2 points pour chaque augmentation d'un écart type du SSÉ. Ainsi, si le SSÉ d'un élève augmente de 0,25 d'un écart type, ses scores en lecture pourraient augmenter de 5 ou 6 points. Une stratégie compensatoire peut également être mise en œuvre à l'échelle de l'école. Toutefois, aucune de ces stratégies n'améliore de façon marquée les résultats des enfants.



Stratégies de redistribution. Une stratégie de redistribution vise l'intégration des enfants marginalisés ou désavantagés dans les écoles régulières. Lorsqu'ils intègrent des écoles régulières, ces enfants profiteraient de l'effet de la composition de l'école dont il a été question précédemment. La fermeture des petites écoles de faible SSÉ et la redistribution des élèves dans d'autres écoles est un exemple de stratégie de redistribution. Une telle stratégie n'a cependant pas une grande incidence sur la réussite et peut même avoir des effets négatifs sur le niveau de scolarité à certains endroits. Par exemple, les enfants risquent de rester plus longtemps à l'école si les écoles près de chez eux offrent tous les niveaux du préprimaire au deuxième cycle du secondaire. Dans les grandes agglomérations, il est possible de diviser les territoires des écoles de manière à réduire la ségrégation attribuable au SSÉ.

Le profil des écoles du Mexique présenté à la figure 12 montre que certaines écoles ont un SSÉ moyen inférieur à -2,0. La **figure 19** montre les effets d'une stratégie de redistribution visant à replacer les élèves des écoles de SSÉ inférieur à -2,0 dans des écoles de SSÉ plus élevé. Toutefois, la majorité des écoles de très faible SSÉ au Mexique sont de petites écoles en milieu rural. Par conséquent, il est peu probable que cette stratégie de redistribution soit viable. De plus, cette stratégie ne contribue que de manière minimale à relever et à maintenir le niveau d'apprentissage.

Figure 18. Effets d'une stratégie compensatoire qui augmente le SSÉ des élèves de faible SSÉ par 0,25 d'un écart type

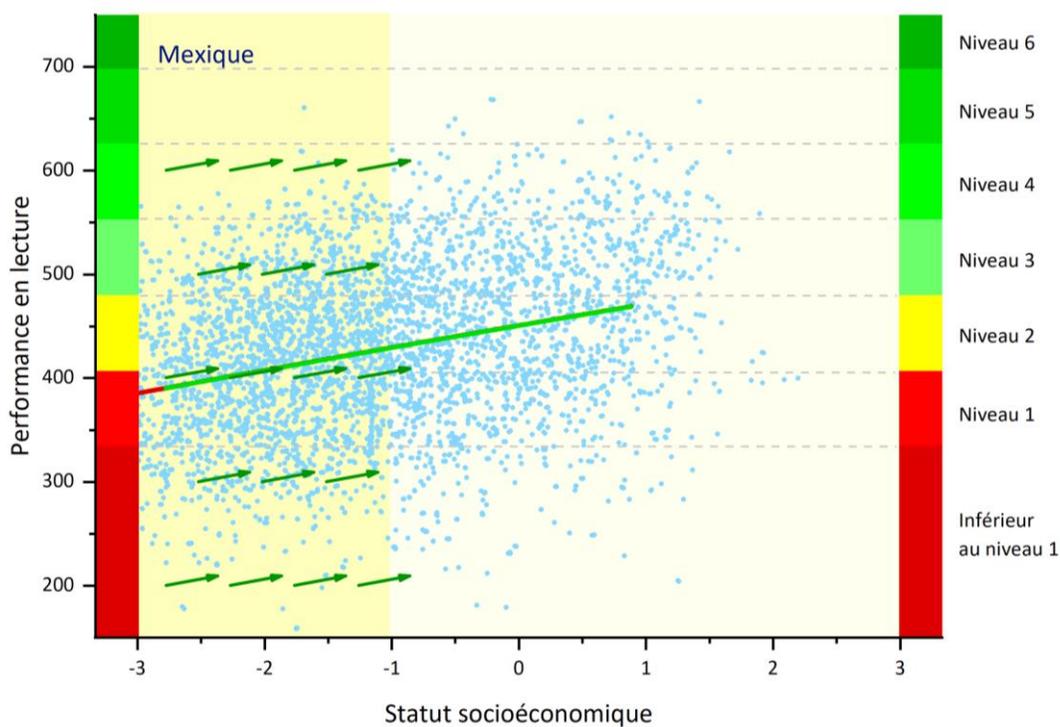
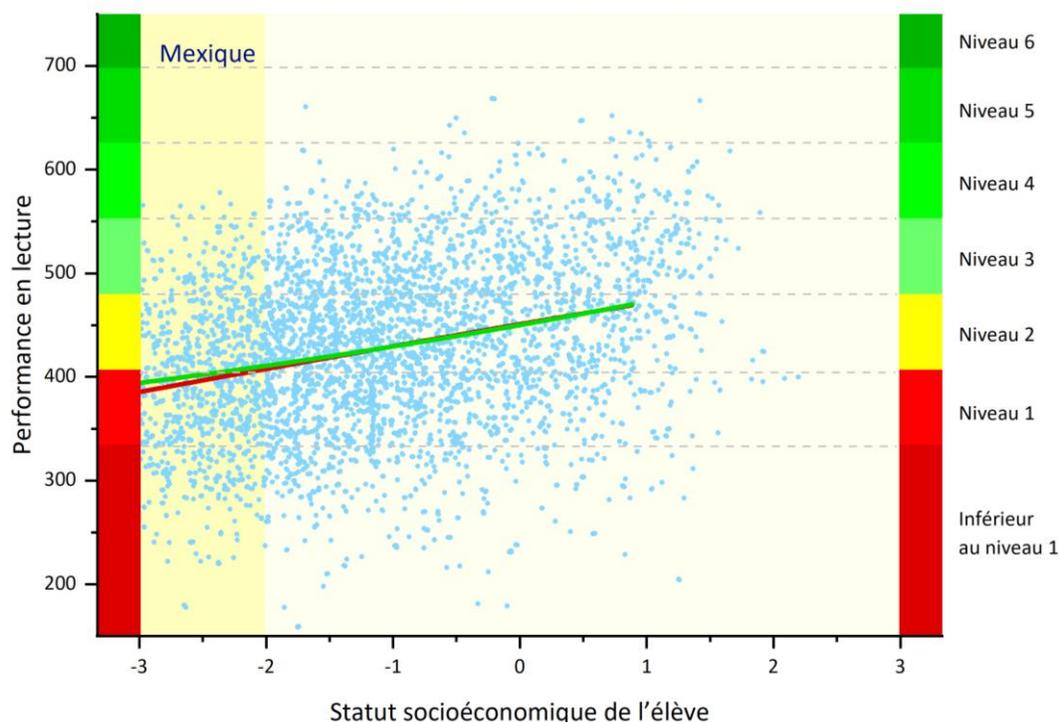




Figure 19. Effets d'une stratégie de redistribution visant à replacer les élèves des écoles de faible SSÉ dans des écoles régulières



Renforcer les bases de la réussite

Les bases de la réussite du modèle de prospérité éducative sont des ensembles de facteurs familiaux, institutionnels et communautaires associés à chacun des six stades de développement. Aux trois derniers stades de développement (pendant lesquels les enfants fréquentent l'école), les facteurs institutionnels considérés comme des bases de la réussite sont les suivants : écoles sûres et inclusives, enseignement de qualité, temps d'apprentissage et ressources matérielles. Lorsque les éducateurs interagissent avec les familles et les leaders des communautés, et les influencent, ils créent un capital social qui contribue au développement des enfants. Par conséquent, dans le présent rapport, un cinquième facteur a été ajouté à l'ensemble de facteurs institutionnels, soit le soutien de la famille et de la communauté.

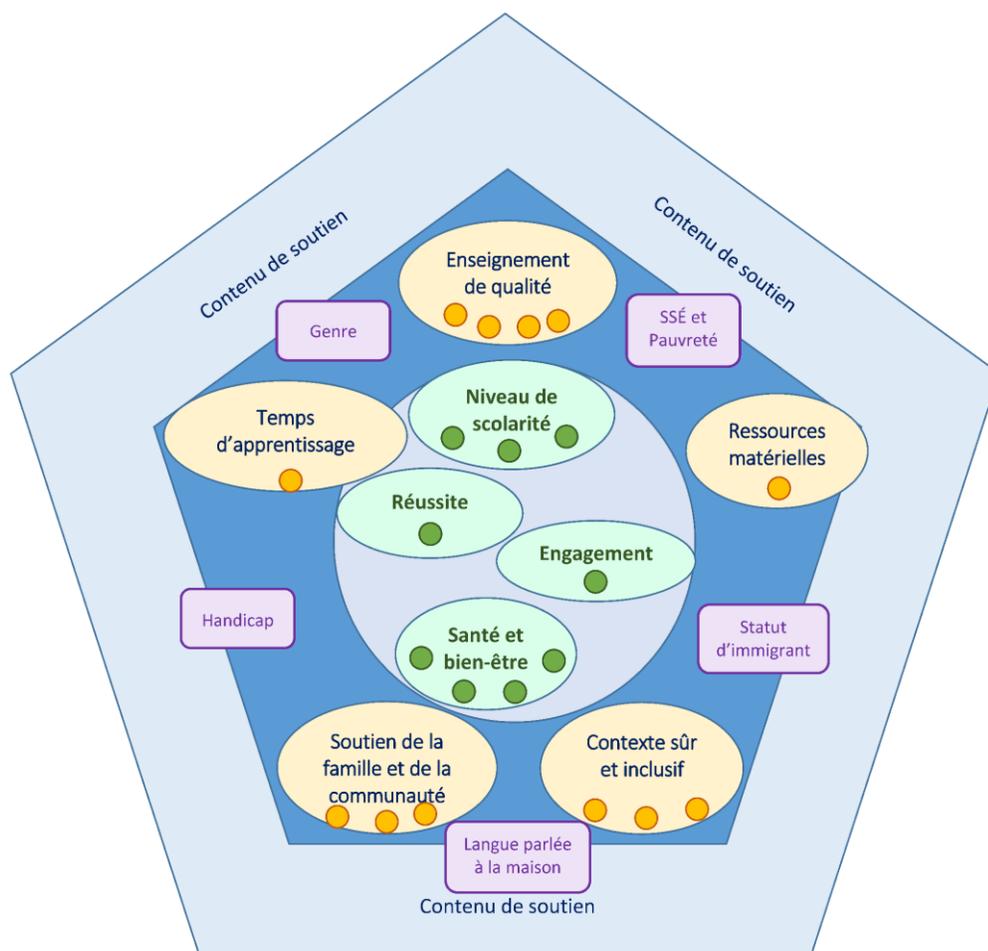
Le modèle de prospérité éducative adopté pour le PISA pour le développement comprend ces cinq bases (OCDE, 2017). La **figure 20** montre les résultats de prospérité, les bases de la réussite et des sous-populations d'intérêt. Elle résume les ensembles de mesures comprises dans les questionnaires des élèves, des enseignants et des écoles. Les quatre résultats de prospérité sont au centre (ovales verts). Il y a une ou plusieurs mesures pour chaque résultat. Par exemple, il y a trois mesures de la réussite scolaire : les compétences en lecture, en mathématiques et en sciences. Il y a également une ou plusieurs mesures associées à chaque base de la réussite (ovales orange). Les sous-populations d'intérêt pour les pays participants sont indiquées dans des rectangles violets. Les questionnaires contextuels comprennent



également d'autres facteurs qui sont pertinents pour les pays participants. Ces facteurs sont le contenu de soutien. L'enquête principale du PISA-D compte trois questionnaires : un pour les élèves, un pour les enseignants et un pour les écoles. Celui des élèves compte 49 questions, celui des enseignants 33 et celui des écoles 28 (OCDE, 2017).

L'objectif est de cibler, pour chaque stade de développement, un petit nombre de mesures critiques sur lesquelles s'entendent tous les intervenants du système scolaire. Ceci pourrait être en quelque sorte le « tableau de bord » du ministre. Les mesures qui sous-tendent chaque base de la réussite fournissent de l'information sur les forces et les faiblesses du système scolaire. Cette information peut être utilisée pour fixer des objectifs à court et à long terme, et élaborer des plans d'action pour les atteindre. Elle peut également jeter les bases d'un plan de communication des orientations de la politique éducative aux divers intervenants à tous les niveaux du système scolaire. Le but d'un tel plan serait d'harmoniser les objectifs et les stratégies à tous les niveaux : au ministère, dans les territoires et dans les écoles.

Figure 20. Cadre de prospérité éducative pour le PISA pour le développement



Contrairement à l'approche s'appuyant sur une logique de cause à effet habituellement adoptée par les recherches sur les effets de l'école, cette approche ne prétend pas étudier tous les facteurs pertinents qui contribuent à la réussite des élèves. Le « leadership des directions » est évidemment un important facteur contribuant à la réussite des élèves, mais ses effets sont modulés par les principales bases de la réussite. Si un programme de formation sur le leadership des directions était mis sur pied par un territoire, ses effets s'exprimeraient selon la capacité d'une direction à assurer un enseignement de qualité dans une école sûre et inclusive. Un tel programme porterait donc sur la définition et la mesure des bases de la réussite, sur les facteurs clés associés à chacune et sur les stratégies à adopter pour poser des bases solides. Cette approche aide également à cerner les stratégies qui risquent de ne produire aucun effet. Souvent, les administrateurs affectionnent des projets qui sont une bonne chose en soi, mais qui ne soutiennent pas les bases menant aux résultats de prospérité. L'objectif de chacun doit être de se consacrer sans relâche aux bases de la réussite.



Il y a une autre différence importante. Les effets des bases de la réussite sur les résultats ont été démontrés par de nombreuses études. Ces effets ne sont pas basés sur des estimations de leurs relations avec les résultats faites dans le cadre d'une récente étude transversale ou de toute autre étude. Les bases de la réussite sont des constructions autonomes représentées par des mesures fiables et valides. Ces mesures peuvent être utilisées pour établir des objectifs à court et à long terme, et élaborer un plan d'action pour les atteindre. Un exemple extrême peut aider à illustrer ce point. La cigarette est mauvaise pour la santé. Nous pourrions avoir un indicateur de santé et de bien-être qui montre la prévalence des jeunes non-fumeurs. Notre préoccupation n'est pas d'estimer les effets de la cigarette sur les mesures de la santé et du bien-être avec le système de suivi. Nous ne voulons pas non plus savoir si le « risque relatif » de la cigarette varie d'un pays à l'autre. Réduire l'usage de la cigarette chez les jeunes est tout simplement la bonne chose à faire.

L'approche de prospérité éducative est conforme aux principes de « cartographie des incidences » (Earl, Carden et Smutylo, 2001). Plus précisément, l'approche de cartographie des incidences « n'est pas fondée sur un système de relations de cause à effet, mais reconnaît plutôt le fait que des événements multiples, non linéaires, mènent au changement. Au lieu d'essayer d'attribuer les incidences à une intervention particulière ou à une série d'interventions, elle examine les liens logiques entre les interventions et la modification du comportement. » (Earl *et al.*, p. 14). Un outil de planification stratégique pourrait consister à élaborer une série de stratégies émanant de chaque base de la réussite. Voici un exemple de question qu'on pourrait poser : « Quelles sont les conditions *nécessaires* et *suffisantes* pour améliorer la qualité de l'éducation? ». On pourrait avancer qu'une condition nécessaire est que les enseignants aient une compréhension approfondie du « modèle simple de la lecture » et un répertoire de stratégies pour enseigner la lecture aux enfants. Cette condition pourrait être liée à un ou plusieurs projets spécifiques. On ne peut toutefois pas attribuer un lien causal à une intervention ou à un projet spécifique parce que d'autres facteurs interviennent également. L'objectif est plutôt de mesurer la « qualité de l'éducation » avec un petit ensemble d'indicateurs clés.

En développant les mesures pour les bases de la réussite, il faut savoir qu'il y a des avantages à les présenter au niveau de l'école. Même si les forces qui agissent sur les résultats varient au sein de l'école et d'une école à l'autre (Rowan, Raudenbush et Kang, 1991), il est plus simple et plus direct de fixer des objectifs pour améliorer les bases de la réussite avec un petit nombre de mesures. Il peut également être avantageux de présenter chacune des mesures sur une échelle de dix points et d'explicitier la règle de notation. Le test décisif est toutefois de s'assurer que les directions des écoles savent comment les bases de la réussite sont mesurées et ce qui doit être fait pour améliorer chacune (p. ex. pour faire passer un score de 4 à 5 ou de 7 à 8 sur l'échelle).

La **figure 21** comprend différents graphiques pour présenter les ressources des écoles du Mexique sur la base des données du PISA 2015. Ces graphiques sont en fait des histogrammes qui résument sur une seule page les indicateurs clés de certaines bases de la réussite. Ces graphiques ne sont que des exemples puisque le PISA 2015 ne comprend qu'un petit nombre de mesures correspondant aux bases de la réussite. Par exemple, on a attribué à chaque école une note pour l'inclusion sur la base d'une mesure du « sentiment d'appartenance » des élèves à leur école. Chaque point sur l'échelle correspond à un pourcentage de 10 %



des élèves ayant un sentiment d'appartenance à leur école. Par exemple, une note de 6 indique qu'entre 50 % et 60 % des élèves de l'école ont un sentiment d'appartenance envers elle. De la même manière, le pourcentage d'élèves dans chaque école qui reçoivent le soutien de leur famille et un temps d'apprentissage suffisant a été établi à partir d'une mesure de la présence positive en classe et à l'école. Les mesures de l'enseignement de qualité et des ressources matérielles sont dérivées de variables au niveau des écoles des ressources en personnel et en matériel provenant des données du PISA 2015.

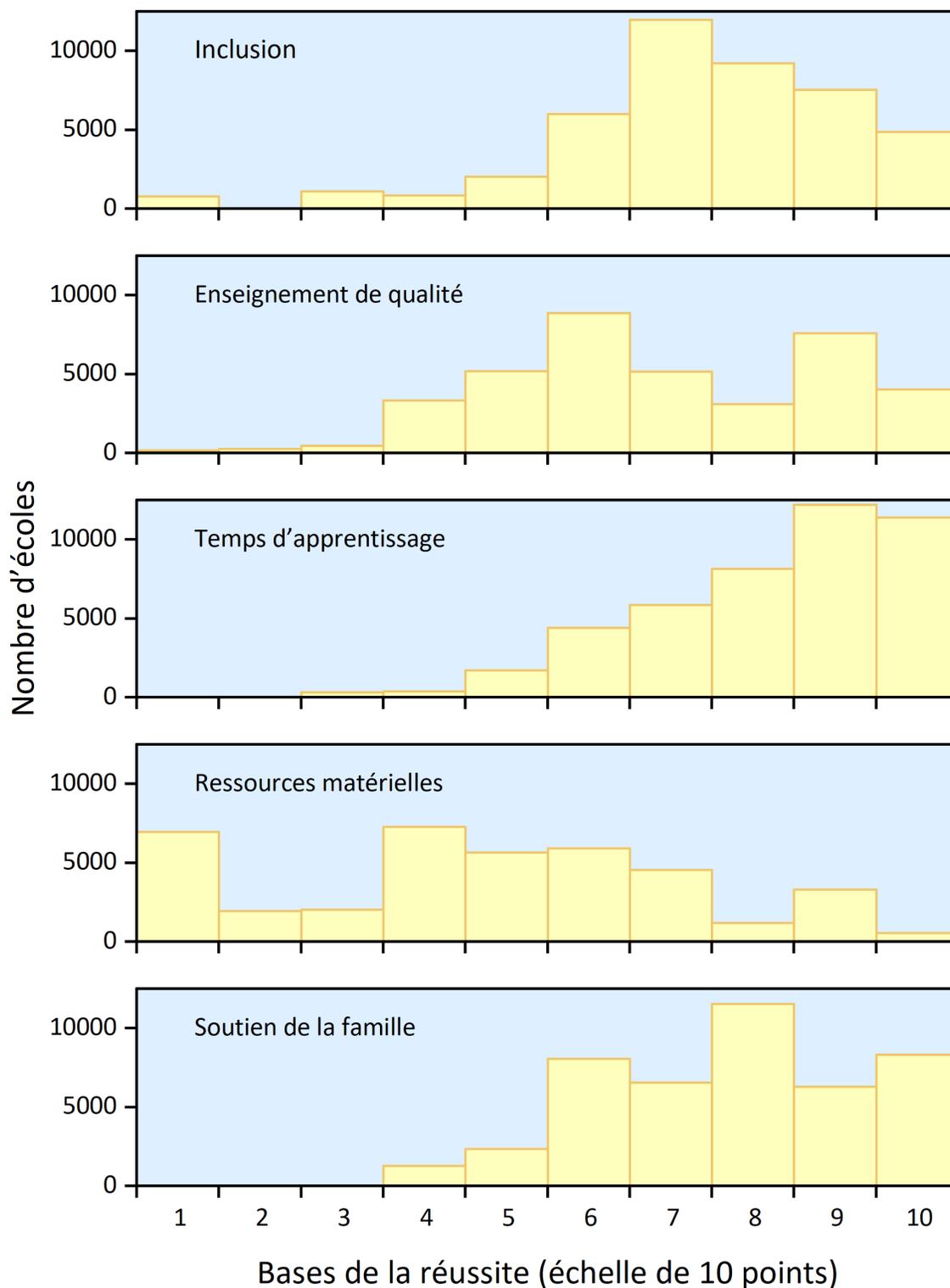
Des analyses de la qualité de l'offre peuvent être réalisées au moyen des techniques décrites plus tôt dans le présent rapport. Pour les mesures au niveau des élèves, les gradients socioéconomiques, les profils des écoles et les graphiques de concentration de la vulnérabilité fournissent de l'information utile pour les politiques d'affectation des ressources. Les profils des écoles fournissent beaucoup d'information pour les mesures au niveau de l'école. Pour certaines sous-populations (p. ex. ethnie, statut d'immigrant), il est possible de construire des variables au niveau de l'école qui dénotent des pourcentages faibles, moyens et élevés, puis de créer des profils comme ceux présentés à titre d'exemples dans les figures 12 et 13.

Dans cet exemple, les graphiques pour le Mexique indiquent qu'il y a une importante variation inter-écoles en ce qui a trait aux ressources et à l'enseignement de qualité. Les mesures du temps d'apprentissage et du soutien des familles quant à elles varient moins et sont plus élevées.⁴ De toute évidence, les données de la principale enquête du PISA ne fournissent pas une évaluation globale des bases de la réussite. De plus, ces données ne fournissent de l'information que sur un échantillon d'écoles tous les trois ans.

⁴ La mesure du PISA pour les ressources matérielles est basée sur les réponses données par des administrateurs scolaires à la question suivante : « La capacité à instruire de votre école est-elle entravée par les problèmes suivants? ». Les choix de réponse à cette question étaient : « pas du tout », « très peu », « dans une certaine mesure » ou « beaucoup? ». Le PISA-D utilise une mesure plus complète basée sur un schéma élaboré par Murillo et Román (2011), lesquels ont interrogé les enseignants sur la disponibilité des ressources didactiques et des services de base, sur leur état et sur l'utilisation qu'ils en faisaient (OCDE, 2017; Tramonte et Willms, 2018).



Figure 21. Ressources des écoles au Mexique



Source : PISA 2015.



Modifier les caractéristiques structurelles des écoles

Dans la majorité des sociétés, l'organisation des écoles et des systèmes scolaires, la matière enseignée et les approches pédagogiques reposent sur de longues traditions. Ce sont les « caractéristiques structurelles » des écoles. Ces caractéristiques structurelles comprennent plusieurs aspects de l'organisation du système scolaire, comme l'âge d'entrée à l'école, le nombre d'années par niveau, la taille des écoles, le programme d'études formel, la langue d'enseignement, les stratégies d'évaluation, la formation professionnelle et les critères de sélection des écoles ou des programmes particuliers. Elles peuvent également faire référence à des aspects de l'organisation de l'école, comme la durée de la journée d'école et des périodes de cours, les modalités d'enseignement et la façon de regrouper les élèves. Les réformes scolaires viennent généralement modifier certaines caractéristiques structurelles de longue date des écoles.

Lorsque les gouvernements entreprennent une réforme visant la restructuration des écoles, ils ont rarement le luxe de pouvoir mener une *véritable* expérience afin de tester une intervention bien définie auprès d'écoles ou d'élèves répartis aléatoirement dans un groupe expérimental ou dans un groupe contrôle. Au contraire, les réformes tendent à se faire lentement avec un passage progressif d'approches plus bureaucratiques, avec des rôles spécialisés et différenciés et des lignes d'autorité formelles du haut vers le bas, à des formes communautaires, avec des rôles variés et une responsabilité partagée pour un ensemble commun d'objectifs (Lee et Smith, 1993; 1995). On peut facilement classer les écoles comme étant soit traditionnelles, soit progressistes et comparer les résultats scolaires et sociaux. Cependant, il est plus difficile d'attribuer une cause à un élément particulier d'une réforme (p. ex. les classes à capacités mixtes, les horaires flexibles ou le fait d'accorder plus de temps aux enseignants pour la planification des cours). Les effets de sélection et les variations dans la fidélité de la mise en œuvre sont trop difficiles à contrôler. De plus, à mesure que la restructuration prend place, des événements critiques ou des changements d'orientation peuvent contrecarrer ou entraver les efforts de réforme, rendant impossible l'évaluation de leurs effets. Les changements de gouvernement et les ingérences des syndicats d'enseignants sont fréquents. De plus, les parents et les leaders des communautés exercent souvent des pressions pour le maintien du *statu quo* (Willms, 2008).

Pour évaluer l'impact des innovations structurelles, il est essentiel d'avoir un solide système de suivi et des mesures fiables et bien définies des résultats de prospérité et des bases de la réussite. Les réformes positives améliorent les résultats et réduisent les inégalités. Elles renforcent les bases. Les données de suivi peuvent être utilisées pour mener de solides quasi-expériences visant à évaluer les effets d'une modification des caractéristiques structurelles des écoles.

Imaginez le cas d'un territoire scolaire convaincu que le décroisement est une structure idéale au primaire, surtout dans les écoles où il y a un fort pourcentage d'élèves vulnérables. Le territoire a en place un solide système de suivi qui lui fournit des données dont il peut se servir pour concevoir et évaluer une intervention. Les administrateurs du territoire décident de lancer une intervention dans le but d'évaluer les effets du décroisement dans un petit nombre d'écoles. Les données de suivi peuvent être mises à profit d'au moins cinq façons.



Premièrement, un territoire doté d'un système de suivi bien conçu peut utiliser ses données comme point de référence dans le cadre d'une étude sur les effets d'une intervention. Par exemple, une intervention en lecture, comme celle du programme *Confident Learners*, pourrait être réalisée dans un petit échantillon d'écoles pour lesquelles il existe des données de référence. Cette étude pourrait prendre la forme d'une quasi-expérience dans le cadre de laquelle l'intervention serait menée dans une région du pays, mais pas dans une autre, ou encore d'une expérience lors de laquelle les écoles seraient réparties aléatoirement dans un groupe expérimental ou dans un groupe contrôle. Le principe fondamental est le suivant : les données de suivi forment un cadre de mesures de référence des résultats de prospérité et des bases de la réussite. Celles-ci permettront de mesurer les changements tout au long du processus de réforme. Il y a un avantage indéniable à procéder de la sorte : comme le territoire n'a pas à consacrer une ou deux années au développement de mesures des résultats ou à l'identification et à la définition de sous-populations, les coûts de démarrage de l'intervention sont moindres. Le territoire peut ainsi consacrer son énergie à surmonter les difficultés inhérentes à la mise en œuvre d'une réforme.

Deuxièmement, les données de suivi peuvent aider à repérer des échantillons stratégiques d'écoles. Les profils des écoles et les graphiques de concentration de la vulnérabilité peuvent être utilisés pour déterminer les écoles qui seront ciblées. Le SSÉ moyen et le pourcentage d'élèves vulnérables de l'école peuvent être utilisés pour définir une strate pour l'échantillonnage aléatoire stratifié (Särndal, Swensson et Wretman, 1992). Par exemple, un échantillon d'écoles potentielles est réparti par strates en fonction du profil des écoles, puis des échantillons aléatoires sont prélevés dans chaque strate.

Troisièmement, s'il est possible de réaliser une évaluation relativement formelle, alors les profils des écoles peuvent être utilisés dans une analyse par bloc aléatoire. Par exemple, des ensembles d'écoles aux profils similaires sont choisis de manière aléatoire pour subir les conditions associées à l'intervention (Matts et Lachin, 1988).

Quatrièmement, les données de suivi peuvent être utilisées après une collecte de données pour renforcer la validité interne de l'évaluation. Les mesures démographiques peuvent servir de contrôles statistiques lors de l'estimation des effets de l'intervention. Elles peuvent également être employées pour évaluer les interactions entre l'intervention et les sujets. Par exemple, les effets des programmes sans niveaux, comparativement à ceux avec niveaux, sont-ils plus importants chez les élèves de faible SSÉ ou chez ceux de SSÉ élevé?

Cinquièmement, le système de suivi fournit une stratégie fiable pour évaluer les changements dans l'égalité des résultats de sous-populations clés.

Enfin, on peut se demander si l'intervention a des effets secondaires souhaitables ou indésirables. Bien que l'objectif principal de l'intervention soit d'améliorer les compétences en lecture des élèves, est-ce que celle-ci fait également en sorte que les élèves sont plus engagés dans leurs apprentissages? Quelle incidence l'intervention a-t-elle sur l'assiduité? Les enseignants sont-ils plus ou moins enclins à utiliser des stratégies d'enseignement efficaces?



V. Faire le suivi de la prospérité éducative

La cadre de prospérité éducative est un puissant outil dont peuvent se servir les ministères de l'Éducation pour suivre un ensemble principal de résultats d'élèves et les forces motrices derrière ceux-ci. Le cadre s'est avéré utile dans les pays à revenu faible, moyen et élevé (Willms, 2018b). Il peut également permettre de suivre les progrès vers l'atteinte des Objectifs de développement durable (ODD) pour 2030. La section qui suit porte sur les stratégies à adopter pour établir des objectifs, présente un exemple de système de suivi complet et aborde les nouvelles orientations que pourrait inspirer le cadre de prospérité éducative en ce qui a trait à l'évaluation de l'éducation.

Établir des objectifs

SMART est un cadre pour l'établissement d'objectifs dans les domaines de la gestion et l'éducation que l'on attribue généralement à Doran (1981). Selon ce cadre, les objectifs doivent satisfaire à cinq critères. Ils doivent être spécifiques, mesurables, attribuables, réalistes et délimités dans le temps. En éducation, les critères « attribuables » et « réalistes » sont habituellement remplacés respectivement par « atteignables » et « pertinents ». Cette version adaptée au domaine de l'éducation est celle à laquelle on fait référence ici.

Spécifiques. Un objectif spécifique énonce clairement ce qui doit être accompli. Les objectifs sont plus susceptibles d'être atteints s'ils sont bien définis et peu nombreux. À des fins de suivi, des objectifs spécifiques peuvent être rédigés pour chacun des résultats de prospérité. Ceux-ci peuvent inclure des cibles pour améliorer la performance scolaire et réduire les inégalités.

À titre d'exemple, voici l'objectif 4.1 du Programme de développement durable à l'horizon 2030 : « D'ici à 2030, faire en sorte que toutes les filles et tous les garçons suivent, sur un pied d'égalité, un cycle complet d'enseignement primaire et secondaire gratuit et de qualité, qui débouche sur un apprentissage véritablement utile. ». Cet objectif devrait être plus spécifique et pourrait, pour ce faire, être divisé en un certain nombre d'objectifs distincts qui définissent des résultats d'apprentissage efficaces. Par ailleurs, le cadre de prospérité éducative fait la distinction entre les résultats et les forces motrices derrière ceux-ci. L'éducation gratuite, équitable et de qualité fait partie des bases de la réussite.

Mesurables. Les objectifs doivent être exprimés en termes absolus et liés à des échelles fixes dans le temps. Par exemple, un pays peut se fixer des objectifs afin de réduire la prévalence des naissances de faible poids. Des critères bien définis sont rattachés à la naissance de faible poids : poids inférieur à 2 500 grammes à la naissance, sans égard à l'âge gestationnel (Organisation mondiale de la santé, 2010). Pour la majorité des variables en éducation, les résultats sont définis par rapport à des échelles standardisées ou sous la forme de variables dichotomiques. L'ampleur de l'effet est une mesure utile si l'échelle est fixe dans le temps. Par exemple : la moyenne et l'écart type sont fixes pour une année de référence et ne sont pas rééchelonnés pour obtenir un nouvel écart type chaque année ou pour chaque cycle d'évaluation. L'échelle du NAPLAN est exemplaire en ce sens qu'elle couvre une vaste tranche d'âge et qu'elle n'est pas rééchelonnée chaque année. Les mesures fondées sur un critère établi sont souvent préférables, comme le pourcentage d'élèves ayant atteint le seuil repère international intermédiaire du PIRLS ou le niveau 2 ou supérieur du PISA. Les



mesures de l'anxiété et de la dépression employées avec les enquêtes notreÉCOLE (The Learning Bar Inc., 2009) (lesquelles sont utilisées par le PISA pour le développement) ont des critères fixes et invariables dans le temps pour les niveaux modérés et graves d'anxiété et de dépression.

Les objectifs énoncés en matière de classement par rapport aux autres pays ou territoires sont inutiles parce qu'un simple petit changement peut modifier considérablement un classement. Par ailleurs, un pays peut occuper un rang plus élevé non pas parce que ses résultats sont meilleurs, mais parce que ceux des autres pays ont reculé.

Atteignables. Faire en sorte que les objectifs soient atteignables est sans doute la plus grande difficulté lorsque l'on établit des objectifs. Ceux-ci doivent être réalistes et atteignables dans une période donnée, tout en constituant un défi pour les éducateurs à tous les niveaux du système scolaire. Comme nous l'avons déjà mentionné, les mesures de l'éducation prennent leur sens lorsqu'elles sont comparées à des normes ou à celles d'autres territoires et lorsqu'elles dessinent des tendances dans le temps. Les évaluations internationales, comme le PIRLS, le PISA et la TIMSS, sont utiles pour établir des objectifs parce qu'elles fournissent un contexte plus large dans lequel évaluer comment se comparent les élèves d'un pays par rapport à ceux d'un autre aux contextes sociaux et économiques comparables. Elles donnent également une idée de l'ampleur des inégalités dans la performance, des iniquités dans l'offre et de la variation des résultats d'une école à l'autre selon le SSÉ.

Chaque école doit recueillir des données sur les résultats de prospérité au niveau de l'élève, à des intervalles réguliers et pour chaque élève. Les résultats du PISA peuvent permettre d'évaluer ce qui peut être fait avec le soutien des autorités nationales et locales, et le talent et la volonté du personnel enseignant. Par exemple, les profils des écoles du Mexique présentés dans les figures 12 et 13 fournissent une indication de la fourchette des scores au PISA et de la prévalence des élèves avec de faibles scores dans toutes les écoles et pour différents niveaux de SSÉ. Si l'on fait abstraction de certaines valeurs aberrantes, les profils suggèrent que la variation inter-écoles des scores du PISA est d'environ 80 à 100 points à tous les niveaux de SSÉ et que la prévalence des enfants vulnérables varie d'environ 40 % à 60 %. Les modèles linéaires hiérarchiques (MLH) à partir desquels les profils sont dérivés fournissent des estimations de la variation des résultats après contrôle du SSÉ moyen de l'école. L'écart interquartiles est une mesure utile pour évaluer si un objectif est atteignable. Pour le Mexique, l'écart interquartiles des scores du PISA est de 36 points et la prévalence des élèves vulnérables est de 20 %. Sur la base de ces données, un objectif atteignable pour une école pourrait être d'augmenter les scores normalisés en lecture d'environ 9 % d'un écart type sur une période de cinq ans et de réduire les niveaux moyens de vulnérabilité de 5 %. Ces cibles correspondent à environ un quart des écarts interquartiles.

Pour suivre les progrès vers l'atteinte de ces objectifs, il faut recueillir des données sur chaque élève à intervalles réguliers, couvrir tous les niveaux de l'école et, idéalement, employer des mesures qui permettent de suivre la progression de chaque enfant par rapport aux résultats de prospérité. Les questionnaires qui recueillent de l'information de base fournissent de l'information sur les bases de la réussite qui peut aider les écoles à suivre leur progression vers la création d'un milieu qui favorise la réussite de tous les élèves. La



plupart des écoles ne peuvent pas faire ceci seules. Le rôle des districts scolaires et des gouvernements locaux et nationaux est donc de fournir aux écoles les outils qui leur permettront de suivre leurs progrès. Dans de nombreux contextes, il faut passer d'un modèle selon lequel les évaluations servent à responsabiliser les écoles à un modèle où les données des tests et des questionnaires sont utilisées comme indicateurs précurseurs pour l'établissement d'objectifs et l'élaboration de plans d'amélioration des écoles.

Le modèle de prospérité éducative incite les administrateurs des ministères de l'Éducation à établir des objectifs sur la base d'une gamme d'évaluations couvrant toute la période allant de la naissance à la fin de l'adolescence. Le modèle encourage également l'élaboration de mesures des bases de la réussite pour chaque stade de développement. Les progrès d'un système scolaire doivent être évalués à partir d'une vaste gamme de mesures, et des objectifs doivent être fixés en conséquence. Pour faire un suivi à ce niveau, il faut réaliser des évaluations régulièrement (chaque année ou au moins tous les deux ans). Ces évaluations peuvent être basées sur des échantillons aléatoires d'écoles et d'élèves.

Les évaluations internationales ne sont pas utiles pour fixer des objectifs parce qu'elles sont trop peu fréquentes et parce que leurs résultats peuvent fluctuer d'un cycle à l'autre et ne pas refléter les changements réels dans la structure d'un système scolaire ou dans l'offre d'une éducation de qualité. Par exemple, les scores en lecture du Mexique au PISA ont reculé de 22 points entre 2000 et 2003, passant de 422 à 400. Les scores en sciences ont également reculé, passant de 422 à 405, tandis que les scores en mathématiques n'ont baissé que de deux points, passant de 387 à 385. Selon Vidal, Díaz et Jarquín (2004), le recul en lecture et en sciences pourrait être attribuable en partie à une hausse du taux de participation au PISA, celui-ci ayant été de 52 % en 2000 et de 58 % en 2003. Cette hausse du taux de participation ne peut cependant pas expliquer les résultats en mathématiques. De plus, le SSÉ moyen de l'échantillon du PISA a régressé légèrement, passant de -1,17 à -1,11. Une autre explication plausible est que les changements sont principalement attribuables à des erreurs d'échantillonnage et de mesure. En 2009, les scores en lecture et en sciences du Mexique étaient revenus à leurs niveaux de 2000, tandis que les scores en mathématiques augmentaient à peu près au même niveau que ceux en lecture et en sciences. Ces données montrent bien que les évaluations internationales ne sont pas utiles pour fixer des objectifs. Si un pays décide d'augmenter la réussite scolaire de 9 % d'un écart type et de réduire le niveau moyen de vulnérabilité de 5 %, les fluctuations attribuables aux erreurs de mesure et d'échantillonnage sont trop importantes pour fournir une mesure fiable de l'évolution de la performance du système. Toutefois, les évaluations internationales peuvent donner de la crédibilité aux résultats des évaluations nationales et leur validité peut être renforcée par des études qui établissent un lien entre les évaluations internationales et nationales (Singer et Braun, 2018).

Pertinents. Le cadre de prospérité éducative énumère quatre ou cinq résultats de prospérité pour chacun des six stades de développement. Leur pertinence n'est pas remise en cause dans le présent rapport. Le critère de « pertinence » est plutôt considéré de deux façons. Premièrement, l'objectif doit être compris et perçu comme important par les intervenants à tous les niveaux du système scolaire, depuis le ministre, son personnel et les administrateurs régionaux jusqu'aux enseignants, parents et élèves. Il est essentiel que tous les intervenants du système scolaire s'entendent sur les mêmes objectifs. Deuxièmement, les objectifs



peuvent être vus comme des moyens de communication. Pour qu'un système scolaire s'améliore, il faut renforcer les bases de la réussite et, pour y parvenir, il faut bien souvent augmenter le financement. Le fait de dévoiler publiquement quelques objectifs peut encourager plusieurs intervenants (p. ex. politiciens, leaders des communautés, donateurs, etc.) à reconnaître leur importance. Les avantages économiques, sanitaires et sociaux qui découlent des résultats de prospérité peuvent renforcer leur « pertinence ». Par exemple, l'amélioration des compétences en littératie est associée à l'augmentation des revenus et des recettes fiscales, et à la réduction de la criminalité, du chômage, de la dépendance à l'aide sociale et des coûts de soins de santé (Hanushek et Woessmann, 2015; Ross et Wu, 1995).

Fréquence. Les études internationales de grande envergure, comme le PIRLS et le PISA, ont des cycles relativement longs. Par ailleurs, les données de ces études ne sont disponibles qu'un an ou plus après leur collecte. Ces évaluations sont utiles pour la planification à long terme. Cependant, pour suivre leurs progrès, les systèmes scolaires ont besoin qu'on leur fournisse des données annuellement, voire plus fréquemment. Ce type de données est particulièrement important pour suivre les compétences en littératie pendant les premières années du primaire.

Un exemple de programme de suivi

La **figure 22** montre un exemple de programme de suivi basé sur le cadre de prospérité éducative. Il s'agit d'une version révisée d'un système d'évaluation conçu pour les populations autochtones au Canada (Willms, 2009b). Les flèches représentent les évaluations qui sont basées sur des recensements et les rectangles celles basées sur des échantillons. Les évaluations suivantes ont permis la mise en œuvre de ce programme dans certains territoires au Canada : Évaluation de la petite enfance à 4 et 5 ans; évaluation *Confident Learners* de la 1^{re} à la 3^e année (flèches orange, à l'exception de l'évaluation de numératie); enquêtes notreÉCOLE pour plusieurs résultats de prospérité et bases de la réussite (flèches violettes). Les données basées sur des échantillons (rectangles orange et rectangles verts) ont été recueillies lors d'évaluations provinciales. Le suivi du niveau de scolarité atteint est fait par les autorités scolaires locales. Le modèle comprend suffisamment de données pour faire le suivi des résultats de prospérité et des bases de la réussite tels qu'ils sont décrits dans le présent rapport.

Nouvelles orientations

Le cadre de prospérité éducative encourage les pays, les territoires et les écoles à adopter de nouvelles méthodes pour faire le suivi des systèmes scolaires. Il les encourage également à utiliser les données autrement pour orienter leur politique éducative. De nombreux pays font passer des tests pour mesurer la performance des élèves à certains niveaux du primaire et du secondaire. Certains territoires recueillent également des données auprès des élèves, des enseignants, des parents et des administrateurs des écoles sur divers aspects « contextuels » des systèmes scolaires. De plus, de nombreux pays participent à des études internationales de grande échelle, comme le PIRLS, la TIMSS et le PISA.

Les compétences en lecture à la fin du parcours secondaire sont le résultat de l'ensemble des expériences d'apprentissage de l'enfant depuis sa naissance et des conditions dans lesquelles s'est déroulée sa période



prénatale. Les résultats d'une évaluation nationale menée en 2017 en Uruguay ont montré que les compétences langagières et cognitives des enfants étaient très variables pendant la petite enfance, soit avant même que ceux-ci n'apprennent à lire dans un cadre formel. Les compétences langagières et cognitives d'au moins un quart des enfants de 5 ans sur le point d'entrer en 1^{re} année avaient une année ou plus de retard par rapport à celles des autres enfants du même âge.

Les enfants qui entrent en 1^{re} année sans avoir les compétences préalables pour apprendre à décoder les mots sont plus à risque d'avoir de faibles compétences en lecture à la fin du premier cycle du primaire. De plus, la majeure partie de la progression en lecture se fait pendant les premières années du primaire. Par exemple, les résultats d'une évaluation nationale des compétences en lecture menée en Australie (NAPLAN) ont montré que les compétences des enfants étaient très variables à la fin de la 3^e année et que le rythme de progression de celles-ci ralentissait ensuite chaque année jusqu'en 9^e année. Après la 9^e année, les résultats du PISA montrent un rythme de progression annuel de seulement 14 % d'un écart type.



Il faut faire davantage de suivi pendant la petite enfance pour deux raisons : parce que les enfants ont des compétences très différentes lorsqu'ils commencent l'école et parce que l'apprentissage de la lecture au primaire est essentiel pour assurer la réussite scolaire à long terme. C'est pourquoi le cadre de prospérité éducative insiste sur l'importance de recueillir des données sur les compétences des enfants avant que ceux-ci entrent à l'école ou peu de temps après, à intervalles réguliers avec les évaluations individuelles réalisées pendant les trois premières années d'école et à l'âge de 8 ou 9 ans, lorsque les enfants devraient faire la transition de « apprendre à lire » à « lire pour apprendre ». Ce changement dans le suivi est nécessaire autant pour les pays à revenu faible, moyen et élevé. Dans leur analyse des résultats scolaires en Afrique du Sud, van der Berg, Spaull, Wills, Gustafsson et Kotzé (2016) insiste sur le fait que « apprendre à lire pour découvrir un sens et avoir du plaisir est la compétence la plus importante que doivent développer les enfants à l'école primaire » (p. 15).

Traditionnellement, les systèmes de suivi se sont surtout intéressés aux compétences en lecture, en mathématiques et en sciences. Le cadre de prospérité éducative quant à lui insiste sur l'importance de suivre une plus vaste gamme de résultats, comme le niveau de scolarité, l'engagement, la santé et le bien-être. Pour les pays à revenu faible et moyen et pour les élèves de zones pauvres des pays à revenu élevé, les mesures du niveau de scolarité atteint doivent enregistrer la fréquentation scolaire et si les élèves ont terminé les niveaux successifs de la scolarité.

Pour le suivi des facteurs contextuels, de nombreux systèmes scolaires se sont concentrés sur l'estimation des différences dans la performance des écoles et ont utilisé les résultats obtenus pour demander aux écoles de rendre des comptes. Bien souvent, les systèmes recueillent des données sur certains facteurs contextuels pour permettre aux analystes d'estimer la « valeur ajoutée » des écoles après ajustement au contexte familial des élèves. Depuis longtemps, on cherche à cerner les facteurs de l'école et du groupe qui produisent une telle valeur ajoutée. Grâce à plus de trente années de recherche sur l'efficacité scolaire, nous savons maintenant quels sont les facteurs qui contribuent à la réussite des élèves. Ces facteurs sont les bases de la réussite du cadre de prospérité éducative. Le cadre de prospérité éducative incite à recueillir des données détaillées sur un petit ensemble de facteurs de base plutôt que sur une vaste gamme de corrélats de la réussite. Les données sur les facteurs de base peuvent être utilisées pour évaluer l'équité de l'offre et fixer un ensemble d'objectifs pour améliorer le système.

Il s'agit de recueillir des « indicateurs précurseurs » pour caractériser le système scolaire plutôt que de recueillir des « indicateurs tardifs », lesquels servent surtout pour la reddition de comptes. Les indicateurs précurseurs peuvent être utilisés au niveau national et régional pour orienter les politiques visant à améliorer les résultats, réduire les inégalités dans les résultats, assurer l'équité de l'offre, affecter les ressources et évaluer les interventions. Ils peuvent également être utilisés au niveau de l'école et du groupe pour guider les pratiques, identifier les élèves vulnérables, établir des objectifs pédagogiques et impliquer les parents d'une manière significative.

Le cadre de prospérité éducative est un modèle simple qui comprend un petit nombre de facteurs de base. Il offre un cadre à partir duquel imaginer une vision commune pour l'amélioration du système qui est



partagée par tous les intervenants, à tous les niveaux : élèves, parents, enseignants, directeurs et administrateurs scolaires. Le système s'améliorera si tous les participants s'appliquent sans relâche à construire de solides bases avec un seul but en tête : donner à tous les enfants la chance de s'épanouir.



Références

- Adlof, S. M., H. W. Catts et J. Lee (2010). « Kindergarten predictors of second versus 8th grade reading comprehension impairments ». *Journal of Learning Disabilities*, 43(4), 332-345.
- Alexander, K. L. et B. K. Eckland (1975). « Contextual effects in the high school attainment process ». *American Sociological Review*, 4, 402-416.
- Alexander, K. L., J. Fennessey, E. L. McDill et R. J. D'Amico (1979). « School SES influences—Composition or context? ». *Sociology of Education*, 52, 222-237.
- Anderson, L. W. (2004). *Increasing Teacher Effectiveness* (2^e éd.). Paris : Institut international de planification de l'éducation de l'UNESCO (IIEP).
- Australian Curriculum, Assessment and Reporting Authority (2017). *NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2017*. Sydney : Australian Curriculum, Assessment and Reporting Authority.
- Backhoff, E., A. Bouzas, E. Hernández et M. García (2007). *Aprendizaje y Desigualdad Social en México: Implicaciones de Política Educativa en el Nivel Básico*. Mexique : Instituto Nacional para la Evaluación de la Educación. ISBN : 968-5924-17-1.
- Banque mondiale (2018a). Données nationales de la Banque mondiale et fichiers de données nationales de l'OCDE. Washington, D. C. : Banque mondiale. Source : <https://data.worldbank.org/indicator/NY.GNP.MKTP.KD>
- Banque mondiale (2018b). *World Development Report 2018: Learning to Realize Education's Promise*. Washington, DC : Banque mondiale. DOI :10.1596/978-1-4648-1096-1
- Barnett, S., S. Ayers et J. Francis (2015). *Comprehensive Measures of Child Outcomes in Early Years: Report to the OECD*. Paris : OCDE.
- Ben-Arieh, A. et I. Frønes (2007). « Indicators of children's well being: What should be measured and why? ». *Social Indicators Research*, 84(3), 249-50.
- Beswick, J.F., E. A. Sloat et J. D. Willms (2008). « Four educational myths that stymie social justice ». *The Educational Forum*, 72(2), 115-128.
- Bhattacharjee, Y. (2015). « Baby brains: The first year ». *National Geographic*, janvier 2015. Source : <http://ngm.nationalgeographic.com/2015/01/baby-brains/bhattacharjee-text>
- Black, R.E., H. A. Lindsay, A. Zulfiqar, L. E. Bhutta, M. Caulfield, E. Majid, Colin Mathers, *et al.* (2008). « Maternal and child undernutrition: Global and regional exposures and health consequences ». *Lancet* 371 (9608) : 243-60.



- Black, M. M., P. Susan, L. C. H. Walker, C. T. Fernald, A. M. Andersen, C. L. D. DiGirolamo, C. McCoy, *et al.* (2017). « Early childhood development coming of age: Science through the life course ». *Lancet*, 389(10064), 77-90.
- Boyce, W. T. et M. S. Kobor (2015). « Development and the epigenome: The 'synapse' of gene-environment interplay ». *Developmental Science*, 18(1), 1-23.
- Boyce, W. T., M. B. Sokolowski et G. E. Robinson (2012). « Toward a new biology of social adversity ». *PNAS*, 109(2), 17143-48.
- Bryk, A. S. et M. E. Driscoll (1988). « The high school community: Contextual influences and consequences for students and teachers ». Madison : National Center on Effective Secondary Schools, University of Wisconsin.
- Caravolas, M., A. Lervåg, S. Defior, S. A. Málková et C. Hulme (2013). « Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies ». *Psychological Science*, 24(8), 1398-1407.
- Center on the Developing Child (2007). *In Brief: The Science of Early Childhood Development*. Cambridge, MA : Center on the Developing Child. Source : <http://developingchild.harvard.edu/resources/inbrief-science-of-ecd>
- Creemers, B. P. M. et L. Kyriakides (2006). « Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model ». *School Effectiveness and School Improvement*, 17, 347-366.
- Cynader, M. S. et B. J. Frost (1999). « Mechanisms of brain development: Neuronal sculpting by the physical and social environment ». Dans D. Keating, C. Hertzman (éds.), *Developmental Health and the Wealth of Nations* (pp. 153-84). New York : Guilford.
- Deary, I. J., S. Strand, P. Smith et C. Fernandes (2007). « Intelligence and educational achievement ». *Intelligence*, 35, 13-21.
- Doran, G. T. (1981). « There's a S.M.A.R.T. way to write management's goals and objectives ». *Management Review. AMA FORUM*. 70(11), 35-36.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Japel, C. (2007). « School readiness and later achievement ». *Developmental Psychology*, 43(6), 1428-1446.
- Earl, S., F. Carden et T. Smutylo (2001). « Outcome mapping: Building learning and reflection into development programs ». Ottawa : Centre de recherches pour le développement international.
- Francis, D. J., S. E. Shaywitz, K. K. Stuebing, B. A. Shaywitz et J. M. Fletcher (1996). « Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis ». *Journal of Educational Psychology*, 88, 3-17.



- Gonzales, E. (2016). *Calculating Standard Errors in International Large-Scale Studies*. Princeton : Educational Testing Service.
- Good, R.H., III, D. C. Simmons et E. J. Kame'enui (2001). « The importance of decision-making utility of a continuum of fluency-based indicators of foundational reading skills for 3rd grade high-stakes outcomes ». *Scientific Studies of Reading*, 5, 257-288.
- Gough, P. B. et W. E. Tunmer (1986). « Decoding, reading and reading disability ». *Remedial and Special Education*, 7(1), 6-10.
- Grek, S. (2009). « Governing by numbers: the PISA 'effect' in Europe ». *Journal of Education Policy*, 24(1), 23-37. DOI : 10.1080/02680930802412669
- Hanushek, E. A. et L. Woessmann (2015). *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge et Londres : MIT Press.
- Hattie, J. (2009). *Visible Learning: A Synthesis of over 800 Meta-Analyses relating to Achievement*. New York : Routledge.
- Hertzman, C. (1999). « The biological embedding of early experiences and its effects on health in adulthood ». *Annals of the New York Academy of Sciences*, 896, 85-85.
- Heugh, K. A. (2013). « Multilingual education policy in South Africa constrained by theoretical and historical disconnections ». *Annual Review of Applied Linguistics*, 33, 215.
- Institut de statistique de l'UNESCO (ISU) (2017). *The Quality Factor: Strengthening National Data to Monitor Sustainable Development Goal 4*. Montréal : Institut de statistique de l'UNESCO.
- Kagan, S. L., E. Moore et S. Bredekamp (éds.). (1995). *Reconsidering Children's Early Learning and Development: Toward Shared Beliefs and Vocabulary*. Washington, DC : National Education Goals Panel.
- Knudsen, E. I. (2004). « Sensitive periods in the development of the brain and behavior ». *Journal of Cognitive Neuroscience*, 16(8), 1412-1425.
- Knudsen, E. I., J. J. Heckman, J. L. Cameron et J. P. Shonkoff (2006). « Economic, neurobiological and behavioral perspectives on building America's future workforce ». *Proceedings of the National Academy of Sciences*, 103(27), 10155-10162.
- Kraemer, H. C., A. E. Kazdin, D. R. Offord, R. C. Kessler, P. S. Jensen et D. J. Kupfer (1997). « Coming to terms with the terms of risk ». *Archives of General Psychiatry*, 54, 337-343.
- Kyriakides, L., C. Christoforou et C. Y. Charalambous (2013). « What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching ». *Teaching and Teacher Education*, 36, 143-152.
- Lee, V. E. et J. B. Smith (1993). « Effects of school restructuring on the achievement and engagement of middle-grade students ». *Sociology of Education*, 66 (July), 164-187.



- Lee, V. E. et J. B. Smith (1995). « Effects of high school restructuring and size on early gains in achievement and engagement ». *Sociology of Education*, 68(4), 241-270.
- Leppänen, U., P. Niemi, K. Aunola et J.-E. Nurmi (2004). « Development of reading skills among preschool and primary school pupils ». *Reading Research Quarterly*, 39, 72-93.
- Levin, H. (2009). « The economic payoff to investing in educational justice ». *Educational Researcher*, 28(1), 5-20.
- Lewin, K. (2015). *Educational Access, Equity and Development: Planning to Make Rights Realities*. Paris : Institut international de planification de l'éducation de l'UNESCO (IIEP).
- Lopez, A. Y. (2016). « Steps Towards a National Implementation of the Early Years Evaluation in Uruguay: Actors and Actions ». Document présenté lors du Inaugural Symposium of the Comparative and International Education Society, Scottsdale, AZ.
- Lorenz, M. O. (1905). « Methods of measuring the concentration of wealth ». Publication de l'*American Statistical Association*, 9(70), 209-219.
- Manyike, T. V. (2012). « A comparison of reading and writing proficiency in home language among Xitsonga speaking learners in South African primary township schools ». *International Journal of Education Science*, 4(2), 143-152.
- Manyike, T. V. (2013). « Bilingual literacy or substantive bilingualism? L1 and L2 reading and writing performance among Grade 7 learners in three township schools Gauteng Province, South Africa ». *Africa Education Review*, 10(2), 187-203. DOI : [10.1080/18146627.2013.812271](https://doi.org/10.1080/18146627.2013.812271)
- Martínez, F. et M. A. Díaz (2016). *México en PISA 2015*. Mexique : Instituto Nacional para la Evaluación de la Educación. Source : <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/316/P1D316.pdf>
- Matts, J. et J. Lachin (1988). « Properties of permuted-block randomization in clinical trials ». *Control Clinical Trials*, 9, 327-344.
- McEwen, B. S. et H. M. Schmeck Jr. (1994). *The Hostage Brain*. New York, NY, É.-U. : Rockefeller University Press.
- McClelland, M. M., F. J. Morrison et D. L. Holmes (2000). « Children at risk for early academic problems: The role of learning-related social skills ». *Early Childhood Research Quarterly*, 15(3), 307-329.
- McClelland, M. M., C. C. Ponitz, E. E. Messersmith et S. Tominey (2010). « Self-regulation: The integration of cognition and emotion ». Dans R. Lerner (Series Ed.) et W. Overton (Vol. Ed.), *Handbook of lifespan human development: vol. 1. Cognition, biology and methods* (pp. 509-553). Hoboken, NJ : Wiley.
- Mislevy, R. J., A. E. Beaton, B. Kaplan et K. M. Sheehan (1992). « Estimation population characteristics from sparse matrix samples of item responses ». *Journal of Educational Measurement*, 29, 133-161.



- Mueller, W. M. et T. L. Parcel (1981). « Measures of socioeconomic status: Alternatives and recommendations ». *Child Development*, 52, 13-30.
- Murillo, F. J. et M. Román (2011). « School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students ». *School Effectiveness and School Improvement*, 22(1), 29-50.
- Nation, K. et M. J. Snowling (2004). « Beyond phonological skills: Broader language skills contribute to the development of reading ». *Journal of Research in Reading*, vol. 27(4), 342-356.
- National Early Literacy Panel (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC : National Institute for Literacy. Source : <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- Nations Unies (2015). *Transformer notre monde : Programme de développement durable à l'horizon 2030*. <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- Nonoyama-Tarumi, Y. et J. D. Willms (2010). « The relative and absolute risks of disadvantaged family background and low levels of school resources on student literacy ». *Economics of Education Review*, 29(2), 214-224.
- OCDE (2001). *Knowledge and Skills for Life – First Results from PISA 2000*. Paris : OCDE.
- OCDE (2017). *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science, Preliminary Version*. Paris : Publications de l'OCDE.
- Organisation mondiale de la santé (OMS) (2010). « Disorders related to short gestation and low birth weight, not elsewhere classified ». *International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) version for 2010*. Source : <http://apps.who.int/classifications/icd10/browse/2010/en#/P07>
- Perfetti, C. A., N. Landi et J. Oakhill (2005). « The acquisition of reading comprehension skill ». Dans M. J. Snowling et C. Hulme (éds.), *The Science of Reading: A Handbook* (pp. 227-247). Oxford, UK : Blackwell.
- Polo, D. S. Z., N. P. A. Araujo et J. C. R. Salceda (2017). « La concepción simple de la lectura en alumnos de 4º de primaria de una escuela fiscal de Quito ». *Alteridad*, 12(1), 228-235.
- Raudenbush, S. W. et J. D. Willms (1995). « The estimation of school effects ». *Journal of Educational and Behavioural Statistics*. 20(4), 307-335.
- Raver, C. C., S. M. Jones, C. Li-Grining, F. Zhai, K. Bub et E. Pressler (2011). « The Chicago School Readiness Project's impact on low income preschoolers' preacademic skills: Self-regulation as a mediating mechanism ». *Child Development*, 82, 362-378. DOI : 10.1111/j.1467-8624.2010.01561.x .
- Rhode Island Kids Count (2005). *Getting Ready: Findings from the National School Readiness Indicators Initiative, a 17 State Partnership*. Providence, RI : Rhode Island Kids Count.



- Riehl, C. J. (2000). « The principal's role in creating inclusive schools for diverse students: A review of normative, empirical and critical literature on the practice of educational administration ». *Review of Educational Research*, 70(1), 55-81.
- Ripoll, J. C., G. Aguado et A. P. Castilla-Earls (2014). « The simple view of reading in elementary school: A systematic review ». *Revista de Logopedia, Foniatría y Audiología*, 34, 17-31.
- Robertson, D. et J. Symons (1996). *Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment*. Londres : London School of Economics, Centre for Economic Performance.
- Rose, J. (2006). *Independent Review of the Teaching of Early Reading*. Londres, HMSO : Department for Education and Skills.
- Rosenshine, B. (2010). *Principles of Instruction*. International Academy of Education. Genève : Bureau international d'éducation de l'UNESCO (BIE).
- Ross, C. E. et C. Wu (1995). « The links between education and health ». *American Sociological Review*, 60 (5), 719-745.
- Rowan, R., S. W. Raudenbush et S. J. Kang (1991). « School climate in secondary schools: A multilevel analysis ». Dans S.W. Raudenbush et J.D. Willms (éds.), *Pupils, Classrooms and Schools: International Studies of Schooling from a Multilevel Perspective*. New York : Academic Press.
- Rust, K. (1985). « Variance estimation for complex estimators in sample surveys ». *Journal of Official Statistics*, 1(4), 381-397.
- Särndal, C. E., B. Swensson et J. Wretman (1992). *Model Assisted Survey Sampling*. New York : Springer-Verlag.
- Scarborough, H. S. (1989). « Prediction of reading disability from familial and individual differences ». *Journal of Educational Psychology*, 81(1), 101-108.
- Scarborough, H. S. (2001). « Connecting early language and literacy to later reading (dis)abilities: Evidence, theory and practice ». Dans S. B. Neuman et D. K. Dickinson (éds.), *Handbook of early literacy research* (pp. 97-110). New York : Guilford Press.
- Schatschneider, C., J. M. Fletcher, D. J. Francis, C. D. Carlson et B. R. Foorman (2004). « Kindergarten prediction of reading skills: A longitudinal comparative analysis ». *Journal of Educational Psychology*, 96(2), 265-282.
- Shonkoff, J. P. et D. A. Phillips (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC : National Academy Press.
- Simola, H. (2005). « The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education ». *Comparative Education*, 41(4), 455-70.



- Singer, J. D. et H. I. Braun (2018). « Testing international education assessments ». *Science*, 360(6384), 38-40. <http://science.sciencemag.org/content/360/6384/38>
- Snow, C. E., M. S. Burns et P. Griffin (1998). *Preventing Reading Difficulties in Young Children*. Washington DC : National Academy Press.
- Storch, S. et G. Whitehurst (2002). « Oral language and code-related precursors to reading: Evidence from a longitudinal structural model ». *Developmental Psychology*, 38(6), 934-947. DOI :10.1037//0012-1649.38.6.934.
- The Learning Bar Inc. (2009). Enquête auprès des élèves notreÉCOLE [instrument de mesure]. Source : <http://www.thelearningbar.com/>
- The Learning Bar Inc. (2011). Évaluation de la petite enfance [instrument de mesure]. Source : <http://www.thelearningbar.com/>
- The Learning Bar Inc. (2016). Confident Learners [instrument de mesure]. Source : <http://www.thelearningbar.com/>
- Torgesen, J., S. Otaiba et M. Grek (2005). « Assessment and instruction for phonemic awareness and word recognition skills ». Dans Catts, H. et Kamhi, A. (éds.), *Reading and Language Disabilities* (pp. 112-145). Boston, MA : Allyn and Bacon.
- Tramonte, L. et J. D. Willms (2018). *New Measures for Comparative Studies of Low- and Middle-Income Countries*. Manuscrit soumis pour publication.
- Trzesniewski, K. H., T. E. Moffitt, A. Caspi, A. Taylor et B. Maughan, B. (2006). « Revisiting the association between reading ability and antisocial behavior: New evidence from a longitudinal behavior genetics study ». *Child Development*, 77, 72-88.
- UNESCO (2005). *Principes directeurs pour l'inclusion : Assurer l'accès à l'Éducation pour tous*. Paris : UNESCO.
- UNICEF (2012). *School Readiness: A Conceptual Framework*. New York, NY : Fonds des Nations Unies pour l'enfance.
- Van der Berg, S., N. Spaull, G. Wills, M. Gustafsson et J. Kotzé (2016). « Identifying binding constraints in education: Synthesis report for the Programme to support Pro-poor Policy development (PsPPd) ». University of Stellenbosch: Department of Economics.
- Vellutino, F. R. et D. M. Scanlon (1987). « Phonological coding, phonological awareness and reading ability: Evidence from a longitudinal and experimental study ». *Merrill-Palmer Quarterly*, 33(3), 321-363.
- Verhoeven, L., J. van Leeuwe et A. Vermeer (2011). « Vocabulary growth and reading development across the school years ». *Scientific Studies of Reading*, 15(1), 8-25.



- Vidal, R., M. A. Díaz et H. Jarquín (2004). *Resultados de las Pruebas PISA 2000 y 2003 en México: Habilidades para la Vida en Estudiantes de 15 Años*. Mexique : Instituto Nacional para la Evaluación de la Educación. Source : <http://publicaciones.inee.edu.mx/buscadorPub/P1/D/202/P1D202.pdf>
- Warwick, L. (2005). « Words to grow on ». *Bulletin of the Centre of Excellence for Early Childhood Development*, 4(1), 2-4.
- Willms, J. D. (1986). « Social class segregation and its relationship to pupils' examination results in Scotland ». *American Sociological Review*, 51, 224-241.
- Willms, J. D. et S. W. Raudenbush (1989). « A longitudinal hierarchical linear model for estimating school effects and their stability ». *Journal of Educational Measurement*, 26(3), 209-232.
- Willms, J. D. et M. A. Somers (2001). « Family, classroom and school effects on children's educational outcomes in Latin America ». *International Journal of School Effectiveness and Improvement*, 12(4), 409-445.
- Willms, J. D. (2011). « Measures of educational equality and equity: A methodological note for the INES Network for the Collection and the Adjudication of System-Level Descriptive Information on Educational Structures, Policies and Practices (NESLI) ». Document présenté à la réunion du NESLI OCDE de mars 2011 au Pays-Bas.
- Willms, J.D. (2003a). *Dix hypothèses sur l'impact des gradients socioéconomiques et des différences communautaires sur le développement de l'enfant*. Ottawa, ON : Direction générale de la recherche appliquée de Développement des ressources humaines Canada.
- Willms, J.D. (2003b). « Literacy proficiency of youth: Evidence of converging socioeconomic gradients ». *International Journal of Educational Research*, 39(3), 247-252.
- Willms, J. D. (2006). *Les écarts en matière d'apprentissage : Dix questions stratégiques sur le rendement et l'équité des écoles et des systèmes d'éducation*. Montréal : Institut de statistique de l'UNESCO.
- Willms, J. D. (2008). « The case for universal French instruction ». *Policy Options*, 29(7), 91-96.
- Willms, J. D. (2009a, octobre). « Classroom diversity and Inclusion: The Educational Advantage ». Présentation plénière à Return to Salamanca – Global Conference on Inclusive Education. Salamanca, Espagne.
- Willms, J. D. (2009b). « Feasibility study for a First Nations assessment system ». Ottawa : Affaires indiennes et du Nord Canada.
- Willms, J. D. (2010). « School composition and contextual effects on student outcomes ».
- Willms, J. D., L. Tramonte, J. Duarte et S. Bos (2012). *Assessing Educational Equality and Equity with Large-Scale Assessment Data: Brazil as a Case Study*. Washington : Inter-American Development Bank.



- Willms, J. D. et L. Tramonte (2018). *The Measurement and Use of Socioeconomic Status in Educational Research*. Manuscrit soumis pour publication.
- Willms, J. D. (2018a). *Educational Prosperity*. Fredericton, NB : The Learning Bar Inc.
- Willms, J. D. (2018b). « Educational Prosperity: An assessment strategy for supporting student learning in low-income countries ». Dans D. A. Wagner, S. Wolf et R. F. Boruch (éds.), *Learning at the Bottom of the Pyramid: Science, Measurement and Policy in Low-Income Countries*. Paris : UNESCO-IIEP.
- Zimmer, R. W. et E. F. Toma (1997). *Peer Effects in Private and Public Schools: A Cross-Country Empirical Analysis*. Lexington : University of Kentucky.



Annexe 1. Taux de croissance de la compétence en lecture, PISA 2000 à 2015

	Point de référence (2000)	(ÉT)	Croissance annuelle	(ÉT)	SSÉ	(ÉT)
Australie	517,1	(2,4)	-1,49	(0,24)	43,8	(1,0)
Autriche	500,1	(3,9)	-1,79	(0,44)	46,0	(1,8)
Belgique	507,2	(3,0)	-0,81	(0,32)	48,2	(1,1)
Canada	520,0	(2,8)	-0,62	(0,25)	33,9	(1,5)
République tchèque	488,5	(4,4)	0,37	(0,39)	49,0	(1,6)
Danemark	491,4	(3,0)	-0,74	(0,27)	37,3	(1,7)
Finlande	548,3	(2,5)	-2,45	(0,30)	31,7	(1,6)
France	521,2	(8,8)	-2,00	(0,84)	50,7	(3,4)
Allemagne	490,0	(5,5)	0,80	(0,70)	45,4	(3,2)
Grèce	478,1	(4,1)	-0,16	(0,31)	34,4	(0,6)
Hongrie	491,4	(5,2)	-0,12	(0,52)	47,9	(2,1)
Islande	487,6	(3,7)	-1,87	(0,33)	26,7	(0,9)
Irlande	524,0	(3,6)	-1,00	(0,40)	38,0	(0,6)
Italie	483,8	(6,2)	0,30	(0,48)	32,1	(1,0)
Japon	504,6	(2,2)	2,20	(0,72)	37,9	(3,6)
Corée	530,1	(7,2)	1,04	(0,89)	32,0	(4,1)
Luxembourg	465,6	(2,9)	1,00	(0,19)	40,4	(1,0)
Mexique	456,5	(37,7)	-1,16	(2,54)	26,6	(13,0)
Pays-Bas	520,4	(6,2)	-1,78	(0,58)	38,9	(1,1)
Norvège	486,1	(6,1)	-0,07	(0,69)	37,3	(1,9)
Nouvelle-Zélande	523,8	(3,0)	-1,10	(0,26)	47,7	(1,5)
Pologne	496,6	(4,0)	2,18	(0,51)	40,2	(1,4)
Portugal	491,3	(2,4)	1,14	(0,19)	31,5	(1,7)
Slovaquie	478,1	(3,6)	-0,99	(0,61)	47,7	(2,3)
Espagne	486,8	(13,0)	0,96	(1,27)	29,5	(1,0)
Suède	506,1	(2,7)	-1,65	(0,51)	39,8	(1,0)
Suisse	504,8	(3,0)	-0,92	(0,39)	42,1	(2,1)
Royaume-Uni	511,6	(11,6)	-1,80	(1,95)	43,0	(2,0)
États-Unis	494,6	(6,7)	-0,20	(0,55)	40,5	(3,5)



Annexe 2. Abrégé des statistiques essentielles en politique éducative

Le calcul des statistiques descriptives dans les études internationales de grande envergure

Poids de rééchantillonnage. Les études internationales emploient deux méthodes tenant compte du plan d'échantillonnage. Le PIRLS et la TIMSS emploient la méthode « jackknife », tandis que le PISA emploie la méthode des répliques répétées équilibrées (*balanced repeated replication*, ou BRR) (Rust, 1985). Par exemple, les données du PISA 2015 sont assorties d'un poids d'échantillonnage global et d'une série de 80 poids BRR. Des statistiques comme la moyenne ou l'écart type peuvent être calculées au moyen du poids d'échantillonnage global. Cependant, les poids BRR sont nécessaires pour une estimation exacte de l'erreur type d'une statistique. Quand une statistique est fondée sur des scores de test, les poids BRR sont adjoints aux valeurs plausibles.

Valeurs plausibles. Dans une étude internationale, les questions sont réparties dans différents cahiers d'examen distribués aléatoirement aux élèves. Le PISA 2015, par exemple, totalisait 810 minutes d'interrogation sur la littératie, les sciences, les mathématiques et la résolution de problème en collaboration; chaque élève devait répondre à un test de 120 minutes. Chaque test était conçu de manière que ses questions couvrent toutes les dimensions de compétence. Les méthodes statistiques employées dans le PISA estiment les distributions des scores des tests, ainsi que la probabilité d'une réponse correcte pour chaque élève, si celui-ci avait répondu à toutes les questions du test (Mislevy, Beaton, Kaplan et Sheehan, 1992). Les données comprennent une série de dix valeurs plausibles pour chaque élève. Ces valeurs sont sélectionnées aléatoirement dans la distribution des compétences mesurées parmi les élèves ayant un profil de réponse et un cadre familial comparables. Pour l'essentiel, cette méthode fournit dix estimations, ou *valeurs plausibles*, indicatrices de la performance de chaque élève s'il avait répondu à la totalité des questions du test.

Des statistiques comme la moyenne, l'écart type ou le coefficient de régression sont calculées de la manière habituelle, au moyen du poids d'échantillonnage global. L'erreur type de la statistique d'intérêt, sans valeurs plausibles, correspond à l'équation suivante :

$$SE_{\varepsilon} = \sqrt{\frac{\sum_{r=1}^R (\varepsilon_r - \varepsilon_0)^2}{R(1 - 0.5)^2}}$$

où ε_0 est la statistique d'intérêt calculée au moyen du poids de sondage complet, ε_r est la statistique d'intérêt calculée au moyen du poids de rééchantillonnage (r), et R est le nombre de répliques, soit 80 dans le cas du PISA.

L'erreur type d'une moyenne calculée avec des valeurs plausibles correspond à l'équation suivante :



$$SE_{\varepsilon} = \sqrt{\left[\sum_{p=1}^P \left(\frac{\sum_{r=1}^R (\varepsilon_{r,p} - \varepsilon_{0,p})^2}{R(1-0.5)^2} \right) * \frac{1}{P} \right] + \left[\left(1 + \frac{1}{P} \right) \frac{\sum_{p=1}^P (\varepsilon_{0,p} - \bar{\varepsilon}_{0,P})^2}{P-1} \right]}$$

où $\varepsilon_{r,p}$ est la statistique d'intérêt calculée avec la valeur plausible p et le poids de rééchantillonnage r , et P est le nombre de valeurs plausibles, soit 10 ans le cas du PISA.

Ces formules sembleront complexes au profane. Cependant, on peut dériver une formule simple à partir d'une matrice de 80 par 1 des poids des répliques pour les statistiques sans valeurs plausibles, ou d'une matrice de 80 par 10 (80 BRR par 10 valeurs plausibles) dans le cas de statistiques avec des valeurs plausibles. Cette méthode peut être employée pour toute statistique, comme l'estimation de l'asymétrie, l'indice de ségrégation ou le coefficient de régression.

Gonzales (2016) décrit la plupart des formules employées dans les études internationales de grande envergure.

[1] Moyenne. La formule de la moyenne est la suivante : $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$, où n est la taille de l'échantillon, et Y_i le score de l'élève i . La moyenne se calcule au moyen du poids d'échantillonnage global, et l'erreur type se calcule au moyen de l'une des deux formules ci-dessus, selon que la moyenne porte ou non sur le score d'un test.

[2] Écart type. La formule de l'écart type est la suivante : $SD = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}$, où n est la taille de l'échantillon, y_i le score de l'élève i , et \bar{y} la moyenne. L'écart type se calcule au moyen du poids d'échantillonnage global, et l'erreur type se calcule au moyen de l'une des deux formules ci-dessus, selon que l'écart type porte ou non sur le score d'un test.

[3] Asymétrie. La formule de l'asymétrie est la suivante : $G = \frac{\sqrt{n(n-1)}}{n-2} \left[\frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\left(\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right)^{\frac{3}{2}}} \right]$

où n est la taille de l'échantillon, y_i le score de l'élève i , et \bar{y} la moyenne. L'asymétrie se calcule au moyen du poids d'échantillonnage global, et l'erreur type se calcule au moyen de l'une des deux formules ci-dessus, selon que l'asymétrie porte ou non sur le score d'un test.

[4] Pourcentage de vulnérabilité. La formule du pourcentage de vulnérabilité est la suivante : $P = \frac{\sum_{i=1}^n D_i}{n}$, où n est la taille de l'échantillon, et D_i est le score dichotomique (0 ou 1) de l'élève i . Le pourcentage de vulnérabilité se calcule au moyen du poids d'échantillonnage global, et l'erreur type se calcule au moyen de l'une des deux formules ci-dessus, selon que le pourcentage de vulnérabilité porte ou non sur le score d'un test.



Gradients socioéconomiques

Les gradients socioéconomiques se calculent au moyen d'un modèle de régression basé sur la méthode des moindres carrés ordinaires qui inclut le SSÉ et le SSÉ², soit :

$$Y_i = \beta_0 + \beta_1 SSÉ_i + \beta_2 SSÉ_i^2 + r_i$$

où Y_i est le score en lecture de l'élève i , et $SSÉ_i$ le SSÉ de l'élève i . Le paramètre r_i correspond au résidu de l'élève i ; c'est-à-dire la déviation du score de l'élève par rapport à la régression linéaire. Les quatre statistiques qui suivent sont dérivées de cette équation.

[5] Niveau du gradient : L'intersection β_0 correspond au score attendu d'un élève ayant un SSÉ de zéro. La variable SSÉ peut être « centrée » sur n'importe quelle valeur, mais on convient généralement de centrer celle-ci sur la moyenne nationale ou, dans le cas des analyses du PISA, sur la moyenne de l'OCDE.

[6] Pente du gradient. Le coefficient β_1 correspond à la pente du gradient socioéconomique. L'hypothèse du gradient est : $H_0: \beta_1 = 0$; $H_1: \beta_1 \neq 0$. La signification statistique de β_1 dépend de la magnitude de l'erreur type, et s'évalue au moyen d'un test t à n-1 degrés de liberté.

[7] Rendement décroissant. Le coefficient β_2 correspond à la pente du SSÉ élevé au carré. Quand β_2 est négatif, les effets sur les résultats diminuent avec l'augmentation du SSÉ. L'hypothèse des rendements décroissants est : $H_0: \beta_2 = 0$; $H_1: \beta_2 \neq 0$. La signification statistique de β_2 dépend de la magnitude de l'erreur type, et s'évalue au moyen d'un test t à n-1 degrés de liberté.

[8] Force du gradient. La force du gradient, R^2 , correspond à la proportion de la variance des résultats attribuable au SSÉ. Il s'agit de la différence entre la variance de Y_i et la variance du résidu, r_i , exprimée comme une fraction de la variance de Y_i .

Pour obtenir l'erreur type de ces statistiques, 80 analyses de régression sont nécessaires (une pour chaque BRR) dans les cas où le résultat est une variable qui n'est pas un score de test. Dans le cas où le résultat est un score de test, l'analyse de régression doit être réalisée à 800 reprises (une pour chaque couple BRR-valeur plausible). Rappelons que ces statistiques sont développées au moyen d'une matrice de 80 par 1 ou de 80 par 10 (voir ci-dessus).

Indices d'inclusion

[9] Indice d'inclusion verticale. L'inclusion verticale correspond à la portion *intra-école* de la variation d'un score :

$$V = \frac{\sigma^2}{\tau}$$

On calcule l'inclusion verticale en appliquant un modèle de régression linéaire hiérarchique « nul » aux résultats :



$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$\beta_{0j} = \gamma_{00} + U_{0j} \quad U_{0j} \sim NID(0, \tau)$$

Dans le modèle de régression hiérarchique, σ^2 est la variance des termes d'erreur du niveau 1, soit la variance *intra-école*, et τ est la variance des termes d'erreur du niveau 2, soit la variance *inter-écoles* :

$$\text{Var}(\varepsilon_{ij}) = \sigma^2 \quad \text{Var}(U_{0j}) = \tau$$

[10] Indice d'inclusion horizontale. L'inclusion horizontale correspond à la portion *intra-école* de la variation du SSÉ :

$$H = \frac{\sigma^2}{\tau}$$

Cet indice se calcule au moyen d'un modèle nul ayant le SSÉ comme variable de résultat.

Modèle de gradient multiniveaux

Le modèle de gradient multiniveaux de base comporte deux niveaux :

$$(1) Y_{ij} = \beta_{0j} + \beta_{1j}SSÉ_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

$$\beta_{1j} = \gamma_{10} + U_{1j} \quad U_{1j} \sim NID(0, \tau_1)$$

où Y_{ij} est le score de l'élève i dans l'école j , $SSÉ_{ij}$ est le SSÉ de l'élève i dans l'école j , et ε_{ij} est le résidu, auquel on attribue une distribution normale, une moyenne de zéro et une variance de σ^2 . Les niveaux des gradients sont β_{0j} , c'est-à-dire les scores espérés des élèves ayant un SSÉ de zéro. C'est pourquoi on les considère comme des moyennes ajustées au SSÉ.

[11] Gradient intra-école. Les pentes des gradients intra-école sont β_{1j} . Elles sont modélisées au niveau 2 comme une pente moyenne, γ_{10} , plus un résidu de la pente moyenne, U_{1j} .

[12] Indice de gradient convergent. La variance des moyennes ajustées au SSÉ, soit les β_{0j} , est la variance de U_{0j} , soit τ_0 . De même, la variance des gradients intra-école, les β_{1j} , est la variance de U_{1j} , soit τ_1 .

Le modèle de régression hiérarchique estime la covariance de U_{0j} et de U_{1j} . La corrélation des moyennes ajustées au SSÉ et des pentes, qui correspond ici à l'indice de gradient convergent, est donnée par l'équation suivante :

$$CGI = \frac{\text{Cov}(U_0, U_1)}{\tau_0 \tau_1}$$



[13] Gradient inter-écoles. Le gradient inter-écoles, qui est une régression des moyennes des résultats des écoles en fonction du SSÉ des écoles, peut également être estimé au moyen d'un modèle à deux niveaux :

$$(1) Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + \gamma_{01} \overline{SSÉ}_{.j} + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

où la pente inter-écoles est γ_{01} .

[14] Effet de la composition sur le SSÉ moyen. L'effet de la composition de l'école sur le SSÉ moyen est estimé en ajoutant le SSÉ moyen des écoles à l'équation des moyennes ajustées des écoles au niveau 2 :

$$(1) Y_{ij} = \beta_{0j} + \beta_{1j} SSÉ_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim NID(0, \sigma^2)$$

$$(2) \beta_{0j} = \gamma_{00} + \gamma_{01} \overline{SSÉ}_{.j} + U_{0j} \quad U_{0j} \sim NID(0, \tau_0)$$

$$\beta_{1j} = \gamma_{10} + U_{1j} \quad U_{1j} \sim NID(0, \tau_1)$$

L'effet de la composition de l'école sur le SSÉ moyen est exprimé par γ_{01} .

[15] Effet de la composition sur la variabilité. L'effet de la composition de l'école sur la variabilité est estimé de la même manière que l'effet sur le SSÉ moyen susmentionné : il s'agit d'ajouter l'écart type du SSÉ à la première équation du niveau 2. L'effet de la composition de l'école sur la variabilité des données est exprimé par γ_{01} .