

JUILLET 2019

RENFORCEMENT DES SYSTÈMES D'ÉVALUATION DE L'APPRENTISSAGE

DOCUMENT DE CONSULTATION POUR LE
MECANISME DE PARTAGE DE CONNAISSANCES
ET D'INNOVATIONS (KIX)

Remerciements

Le présent document a été préparé par Kate Anderson, consultante indépendante et fondatrice de la société Unbounded Associates, sous la direction de Ramya Vivekanandan, spécialiste senior en éducation et responsable-thématique des systèmes d'évaluation de l'apprentissage au Secrétariat du PME.

Nous remercions sincèrement les personnes suivantes qui ont examiné le présent document et nous ont apporté leur précieux concours : Hassane Djibo Adamou, Ray Adams, Michelle Belisle, Jean-Marc Bernard, Chris Berry, Suman Bhattacharjea, Moritz Bilagher, Ryan Burgess, Manuel Cardoso, Esther Care, Talia de Chaisemartin, Marguerite Clarke, Luis Crouch, Dirk Hastedt, Sarah Hennell, Hilaire Hounkpodoté, Baela Raza Jamil, Luke Justice, Anil Kanjee, Joseph François Désiré Kauphy, Janeanne Kiviu, Asefa Leta, Silvia Montoya, Karen Mundy, Abdulkadir Yusuf Nuh, Eileen O'Malley, Folake Olatunji-David, Atilio Pizarro, Caroline Pontefract, Sara Ruto, Mioko Saito, François Sawadogo, Ursula Schwantner, Ernesto Trevino, Dan Wagner, Michael Ward et Camilla Woeldike.

Nous souhaitons aussi remercier l'UNESCO Dakar (Secrétariat du Réseau des éducateurs pour la transformation – TALENT) et l'UNESCO Bangkok (Secrétariat du Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique – NEQMAP) pour leur aide dans l'organisation des sessions de consultation qui nous ont permis de recueillir des retours d'information sur le présent document à l'occasion des ateliers régionaux TALENT et NEQMAP, tenus respectivement à Dar es Salam (Tanzanie), en juillet 2018, et à Pénang (Malaisie), en septembre 2018. Ces consultations rassemblaient des représentants des pays suivants : Australie, Bangladesh, Bénin, Bhoutan, Burkina Faso, Burundi, Cambodge, Comores, Éthiopie, Fidji, Gambie, Iran, Kirghizistan, Lesotho, Libéria, Malawi, Malaisie, Maldives, Mali, Mongolie, Myanmar, Népal, Niger, Ouganda, Ouzbékistan, Papouasie-Nouvelle-Guinée, Philippines, République centrafricaine, RDP lao, Sénégal, Soudan du Sud, Sri Lanka, Tanzanie, Tchad, Thaïlande, Timor Leste, Viet Nam et Zimbabwe.

Table des matières

Résumé analytique.....	6
1. Introduction	10
2. Préparation du document et processus de consultation	12
3. Obstacles à la mise en place de systèmes solides d'évaluation de l'apprentissage	13
4. Appui du PME aux systèmes d'évaluation de l'apprentissage.....	16
5. Biens mondiaux et innovations relatifs à l'évaluation de l'apprentissage.....	18
6. Carences des biens mondiaux disponibles.....	27
7. Domaines d'investissement envisageables.....	28
Références	36
Annexe A. Fiche d'information du PME sur l'évaluation de l'apprentissage	40
Annexe B. Évaluations nationales et internationales réalisées dans les PDP, de 2010 à 2017	43

Note sur le mécanisme KIX

Le financement thématique du mécanisme de partage de connaissances et d'innovations (KIX) du partenariat mondial pour l'éducation (PME) appuiera les initiatives régionales et mondiales qui utilisent l'échange de connaissances, les données probantes et les innovations pour aider les pays en développement à relever les principaux défis du secteur de l'éducation. Ce processus appuiera :

- le renforcement des capacités et l'échange de connaissances entre pays en développement : activités renforçant les capacités nationales grâce à l'échange et à la vérification mutuelle ; création de modules d'apprentissage et d'outils de diagnostic ; les rencontres en personne ;
- le recueil de données probantes et les évaluations : activités visant à consolider ou enrichir les connaissances afin d'améliorer les résultats éducatifs et les systèmes nationaux d'éducation ;
- les projets pilotes en matière d'innovations : pilotage d'approches, de méthodes, d'outils ou de produits qui résolvent des défis persistants en matière d'éducation.

Les investissements seront guidés par les priorités des pays en développement, et affectés selon un processus compétitif par un agent partenaire indépendant. Les produits du savoir, les projets pilotes en matière d'innovation et les outils connexes élaborés grâce au financement du mécanisme KIX seront partagés dans le cadre de la plateforme d'échange de connaissances afin de favoriser leur adoption.

Le présent document a pour objet de décrire les efforts actuels de renforcement des systèmes d'évaluation de l'apprentissage et de stimuler la discussion et les débats autour des domaines potentiels dans lesquels le KIX pourrait investir. Il s'inscrit dans une série de documents de travail, rédigés pour soutenir l'engagement et la consultation des pays en développement partenaires (PDP) et des experts techniques dans la conception initiale du mécanisme KIX du PME. Les idées présentées dans la version initiale de ce document ont servi de point de départ aux débats et ont évolué sensiblement au fil du processus de consultation pour aboutir à la présente version actualisée.

Sigles et abréviations

A4L	Évaluation au service des apprentissages
ADEA	Association pour le développement de l'éducation en Afrique
AEEA	Association pour l'évaluation éducative en Afrique
AIEE	Association internationale pour l'évaluation éducative
AMR	Activités mondiales et régionales
ANLAS	Analyse des systèmes nationaux d'évaluation des acquis
ASER	Rapport annuel sur la situation de l'éducation
CLA	Évaluations citoyennes
EGMA	Évaluation de l'apprentissage précoce des mathématiques
EGRA	Évaluation de l'apprentissage précoce de la lecture
ESPIG	Financement pour la mise en œuvre du programme sectoriel de l'éducation
GAML	Alliance mondiale pour le suivi de l'apprentissage
GP-LA	Bonnes pratiques en matière d'évaluation de l'apprentissage
IEA	Association internationale pour l'évaluation du rendement scolaire
IIEP	Institut international de planification de l'éducation
ISU	Institut de statistique de l'UNESCO
KIX	Partage de connaissances et d'innovations
LACI	Indice de la capacité à évaluer l'apprentissage
LaNA	Évaluation des aptitudes à la lecture, à l'écriture et au calcul
LLECE	Laboratoire latino-américain pour l'évaluation de la qualité de l'éducation
MICS	Enquête par grappes à indicateurs multiples
NALA	Réseau pour l'évaluation de l'apprentissage en Afrique
NCME	National Council on Measurement in Education
NEQMAP	Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique
OCDE	Organisation de coopération et de développement économiques
ODD	Objectifs de développement durable
ONG	Organisation non gouvernementale
PASEC	Programme d'analyse des systèmes éducatifs de la CONFEMEN
PDP	Pays en développement partenaire
PEID	Petit État insulaire en développement
PFC	Pays fragile ou touché par un conflit
PILNA	Évaluation régionale des aptitudes à la lecture, à l'écriture et au calcul de la communauté du Pacifique
PIRLS	Programme international de recherche en lecture scolaire
PISA	Programme international de l'OCDE pour le suivi des acquis des élèves
PISA-D	Initiative PISA pour le développement
PME	Partenariat mondial pour l'éducation
SABER	Approche systémique pour de meilleurs résultats dans le domaine de l'éducation
SACMEQ	Consortium d'Afrique australe et orientale pour le suivi de la qualité de l'éducation
SEA-PLM	Initiative pour la mesure de l'apprentissage à l'école primaire en Asie du Sud-est,
SIGE	Système d'information pour la gestion de l'éducation
TALENT	Enseignement et apprentissage : Réseau des éducateurs pour la transformation
TERCE	Troisième étude explicative et comparative régionale
TIMSS	Étude internationale sur les tendances de l'enseignement des sciences et des mathématiques
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture

Termes clés

Évaluation de l'apprentissage. Processus consistant à rassembler et à évaluer les informations sur ce que les élèves connaissent, comprennent et peuvent faire.

Système d'évaluation de l'apprentissage. Ensemble de politiques, d'institutions, de pratiques et d'outils d'évaluation destinés à recueillir et à utiliser des informations sur les acquis des élèves dans le système d'éducation.

Évaluation à l'échelle de la classe. Ensemble de méthodes utilisées pour obtenir des données probantes sur les connaissances, les compétences et les comportements des élèves qui serviront aux enseignants pour améliorer l'apprentissage et l'enseignement.

Examens. Évaluations réalisées pour mesurer chez les élèves l'acquisition des connaissances et des compétences, principalement aux fins de la sélection.

Évaluations à grande échelle. Évaluations réalisées au niveau des systèmes servant au suivi et à la production d'informations, pour les décideurs et les praticiens, sur les niveaux globaux de performance, sur l'évolution de ces niveaux au fil du temps, et sur les facteurs qui les déterminent.

Sources : Clarke (2012), Wagner (2012).

Résumé analytique

Importance des systèmes d'évaluation de l'apprentissage

L'utilisation efficace des données d'évaluation est une composante essentielle du système d'éducation. Il est en effet impossible de savoir si un pays offre une éducation de qualité à l'ensemble de ses citoyens sans disposer de données sur le processus d'apprentissage des élèves. L'évaluation de l'apprentissage peut être un outil efficace pour améliorer les systèmes d'éducation nationale lorsqu'elle s'appuie sur des données contextuelles. Un système d'évaluation de l'apprentissage englobe les politiques, les institutions, les outils et les pratiques d'évaluation qui permettent de recueillir et d'utiliser des informations sur les acquis des élèves dans le système d'éducation¹. Il peut comprendre des évaluations conçues par les enseignants pour étayer leur pratique, des examens nationaux administrés à tous les élèves, et des évaluations basées sur des échantillons, utilisées principalement par les responsables publics. Un système rigoureux d'évaluation des apprentissages peut aider les élèves, leurs parents, les enseignants et les administrateurs d'établissements scolaires à suivre les progrès réalisés et à intervenir pour aider ceux qui ont des difficultés d'apprentissage.

L'évaluation des résultats de l'apprentissage est au cœur du programme mondial d'éducation défini dans les objectifs de développement durable (ODD) et par le cadre d'action Éducation 2030 ; elle constitue donc une priorité mondiale². Le quatrième objectif de développement durable énonce un ensemble ambitieux de cibles, et notamment l'amélioration des résultats d'apprentissage pour tous, de la petite enfance à l'âge adulte. La moitié des dix objectifs liés à l'éducation de l'ODD 4 ont trait à l'apprentissage et aux compétences³. Malgré les bénéfices escomptés d'un suivi des résultats

¹ Clarke, Marguerite. « *What Matters Most for Student Assessment Systems: A Framework Paper* ». Washington, DC : Banque mondiale, 2012.

² ODD 4 : « Garantir une éducation de qualité sans exclusion et équitable et promouvoir des possibilités d'apprentissage tout au long de la vie pour tous ».

³ Les cibles 4.1, 4.2, 4.4, 4.6 et 4.7 liées aux ODD et leurs indicateurs correspondants sont liées aux résultats d'apprentissage.

d'apprentissage, de nombreux pays ne disposent pas de systèmes d'évaluation de l'apprentissage appropriés et de qualité. C'est spécialement vrai dans les pays à faible revenu, ceux touchés par la fragilité ou les conflits et les petits États insulaires en développement (PEID) dont de nombreux pays en développement partenaires du PME font partie.

Obstacles à la mise en place de systèmes solides d'évaluation de l'apprentissage

Un examen sur dossier et des consultations menées auprès des PDP et d'experts internationaux ont mis en lumière quatre grands obstacles à la mise en place de systèmes solides d'évaluation de l'apprentissage : 1) qualité insuffisante des outils d'évaluation et manque de compétences techniques pour la préparation et la réalisation des évaluations et l'analyse de leurs résultats ; 2) mauvais alignement des divers types d'évaluations utilisées, et intégration déficiente au sein des systèmes nationaux d'éducation ; 3) exclusion fréquente des enfants les plus marginalisés, handicapés ou non scolarisés ; 4) mauvaise utilisation des résultats des évaluations de l'apprentissage découlant notamment de la façon dont ces évaluations sont conçues et distribuées, et absence des personnes qualifiées et des ressources requises pour procéder aux changements nécessaires.

Appui du PME aux systèmes d'évaluation de l'apprentissage

La mise en place de systèmes solides d'évaluation de l'apprentissage compte parmi les priorités du PME, comme en attestent son cadre de résultats et son modèle de financement. Améliorer les résultats d'apprentissage et les rendre plus équitables est le premier des trois buts énoncés dans le Plan stratégique 2020 du PME. Le modèle de financement du PME exige par ailleurs des pays qui présentent une requête de financement pour la mise en œuvre de leur programme d'éducation qu'ils disposent d'un système de suivi des acquis scolaires ou d'un plan pour la mise en place d'un tel système. En 2017, le Conseil d'administration du PME a décidé que si les financements disponibles ne suffisaient pas à créer de tels systèmes, il conviendrait d'utiliser à cette fin les ressources financières du PME

Le cadre de résultats du PME permet de suivre la proportion des PDP qui auront réussi à améliorer les acquis scolaires de leurs élèves dans le secteur de l'éducation de base de 2016 à 2020, et de déterminer si leurs systèmes d'évaluation de l'apprentissage répondent aux normes de qualité. Les financements du PME pour la mise en œuvre de programmes nationaux aident actuellement 29 pays (71 % des financements en cours) à élaborer différents aspects de leurs systèmes d'évaluation de l'apprentissage. Les bénéficiaires des financements du PME doivent, pour accéder à ces derniers, démontrer qu'ils disposent de systèmes de données de bonne qualité ou d'une stratégie pour les mettre en place. En 2016, seuls 32 % des PDP disposaient d'un tel système ou d'une stratégie pour sa mise en place.

Pour aider les PDP à atteindre ces ambitieux objectifs, le PME appuie également les activités régionales et internationales qui renforcent les connaissances, les données probantes et les capacités nationales. Les financements précédents ont aidé l'Institut de statistique de l'UNESCO à définir les bonnes pratiques en matière d'évaluation des compétences en lecture dans les petites classes et d'inventaire des évaluations d'apprentissage. Ils ont aussi aidé le Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique (NEQMAP) à recueillir des données probantes sur le plan régional et à renforcer les capacités dans ce domaine. L'initiative Évaluation au service des apprentissages (A4L) a été lancée en juillet 2017, à la suite de ces activités, pour renforcer la capacité des systèmes nationaux d'évaluation de l'apprentissage. Elle plante le décor pour le mécanisme KIX en élaborant un outil qui aidera les pays à analyser leurs systèmes d'évaluation de l'apprentissage et contribuera ainsi aux efforts de renforcement de ces systèmes. L'Initiative A4L apporte aussi une assistance aux réseaux régionaux qui aident les pays à mettre en place les systèmes d'évaluation répondant aux besoins de leurs programmes nationaux et à leurs priorités stratégiques. Les financements KIX renforceront ces efforts en favorisant encore

davantage les échanges de connaissances et de nouveaux biens mondiaux et en s'appuyant sur ceux qui existent déjà. Ces financements sont décrits à la section suivante.

Biens mondiaux et innovations relatifs à l'évaluation de l'apprentissage

Les pays qui souhaitent évaluer l'apprentissage disposent d'une variété d'outils, d'études et de recherches. Les évaluations régionales et internationales comparatives offrent des données transnationales précieuses sur l'apprentissage des élèves dans plusieurs pays et favorisent souvent le renforcement des capacités et les occasions d'apprentissage par les pairs⁴. Les évaluations des compétences relatives aux fondamentaux de la lecture et du calcul — par exemple, évaluations citoyennes et évaluations de l'apprentissage précoce de la lecture (EGRA) — ne sont pas conçues pour être comparées d'un pays à l'autre. Cependant la communauté des utilisateurs de ces évaluations peut aider les pays à en faire bon usage. Un certain nombre de plateformes en ligne et d'organisations professionnelles facilitent l'échange de connaissances, en particulier pour ceux travaillant sur les évaluations à grande échelle⁵.

D'autres initiatives soutiennent le renforcement des capacités, notamment des livres, des guides et des cours élaborés pour les pays à revenu faible et intermédiaire. Par exemple, l'Alliance mondiale pour le suivi de l'apprentissage (GAML) hébergée par l'Institut de statistique de l'UNESCO (ISU), a mis au point des directives sur les bonnes pratiques en matière d'évaluation de l'apprentissage et plusieurs petits guides sur l'évaluation de l'apprentissage destinés aux personnes mettant en œuvre les évaluations⁶. Le module relatif à l'évaluation des acquis des élèves de l'Approche systémique pour de meilleurs résultats dans le domaine de l'éducation (SABER) de la Banque mondiale offre des rubriques fondées sur des faits probants pour l'évaluation du système d'évaluation de l'apprentissage d'un pays. Ce module dispose de données sur 38 pays. Le PME s'emploie enfin à mettre au point un outil complémentaire d'aide ou diagnostic des systèmes d'évaluation de l'apprentissage connu sous le nom d'Analyse des systèmes nationaux d'évaluation des acquis (ANLAS).

Parmi les biens mondiaux existants figurent enfin des initiatives novatrices en matière d'évaluation de l'apprentissage et d'utilisation des résultats pour l'amélioration de l'apprentissage. L'ISU s'emploie à améliorer sa méthode de soutien social pour mettre en équation les compétences des élèves en lecture et en calcul mesurées dans divers pays et à l'aide de divers outils d'évaluation. Les ONG ont également recours à des méthodes novatrices d'évaluation de l'apprentissage pour un enseignement adapté aux

⁴ Notamment le module sur les compétences fondamentales d'apprentissage des enquêtes par grappes à indicateurs multiples (MICS) de l'UNICEF, le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC), l'Évaluation régionale des aptitudes à la lecture, à l'écriture et au calcul de la communauté du Pacifique (PILNA), le Programme international de recherche en lecture scolaire (PIRLS), le Programme international pour le suivi des acquis des élèves (PISA et PISA-D), l'Initiative pour la mesure de l'apprentissage à l'école primaire en Asie du Sud-est (SEA-PLM), le Consortium d'Afrique australe pour le suivi de la qualité de l'éducation (SACMEQ), la Troisième étude explicative et comparative régionale (TERCE), et l'Étude internationale sur les tendances de l'enseignement des sciences et des mathématiques (TIMSS).

⁵ Y compris la plupart des agences d'évaluation transnationale énumérées dans la note 4, ainsi que le Portail mondial de l'IPE (UNESCO), l'Alliance mondiale pour le suivi de l'apprentissage (GAML), l'Association internationale pour l'évaluation éducative (AIEE), le National Council on Measurement in Education (NCME), le Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique (NEQMAP), Enseignement et apprentissage : Réseau des éducateurs pour la transformation (TALENT), le Réseau pour l'évaluation de l'apprentissage en Afrique (NALA) et l'Association pour l'évaluation éducative en Afrique (AEEA).

⁶ ISU (Institut de statistique de l'UNESCO) et ACER (Australian Council for Educational Research). « *Principles of Good Practice in Learning Assessment* ». Camberwell, Victoria: ACER, 2017.

capacités des enfants⁷ et pour la mise au point d'évaluations en classe mesurant les compétences du XXI^e siècle⁸. Diverses compétences comme la résolution de problèmes et le sens citoyen font aussi de plus en plus l'objet d'évaluations approfondies dans le cadre des évaluations transnationales.

Disparités des biens mondiaux disponibles

La plupart des biens mondiaux existants sont axés sur les points suivants : 1) appui aux groupes de pays pour la conception et l'administration d'évaluations particulières, à grande échelle, de l'apprentissage ; 2) regroupement des réseaux de professionnels pour faciliter l'échange de connaissances sur les enjeux de l'évaluation ; 3) ressources en ligne comprenant des directives générales sur la mise au point de méthodes d'évaluations de l'apprentissage de qualité.

Les PDP ont souligné, à une écrasante majorité, qu'il existe des disparités dans les directives sur la formation des enseignants à l'administration et à l'utilisation des évaluations en classe, et dans les bonnes pratiques de réforme des examens. Bien que la participation aux évaluations à grande échelle assure dans une certaine mesure le renforcement des capacités de certains représentants ministériels, les PDP ont fait remarquer que ces directives et bonnes pratiques ont besoin d'être élargies pour instaurer une culture globale d'évaluation. Ils ont enfin attiré l'attention sur les disparités des systèmes d'évaluation de l'apprentissage, et demandé des conseils sur la façon de créer de solides capacités institutionnelles pour assurer le suivi de l'ensemble des portefeuilles nationaux de systèmes d'évaluation. Tant les experts internationaux que les PDP ont noté l'absence de données comparables sur les acquis scolaires au niveau international ; ils ont exprimé le souhait que plus de pays participent aux évaluations internationales.

Domaines d'investissement envisageables

Les consultations menées ont débouché sur quatre possibilités, pour les biens mondiaux financés par le mécanisme KIX, de jouer un rôle crucial. Ces possibilités s'harmonisent également avec le cadre de résultats et le modèle de financement du PME:

Biens mondiaux à l'appui des institutions nationales d'évaluation de l'apprentissage, y compris les ministères de l'éducation et les agences autonomes ou semi-autonomes d'évaluation. Cet appui pourrait porter sur des activités de renforcement des capacités visant à aider les pays à établir des cadres nationaux d'évaluation et des protocoles d'échange régional axés sur la création de solides systèmes d'évaluation. Il pourrait aussi inclure l'élargissement de la portée des outils de diagnostic des systèmes d'évaluation de l'apprentissage et des améliorations recommandées, l'examen des succès remportés en matière de réforme des examens publics et des évaluations fondées sur les compétences, et l'aide à l'innovation en faveur des différents groupes de pays afin d'élaborer de concert des systèmes d'évaluation des compétences au-delà de la seule acquisition de la lecture et du calcul.

Biens mondiaux pour le renforcement des compétences en matière d'évaluation de l'apprentissage à l'échelle du système d'éducation, des enseignants aux dirigeants ministériels. Cet appui pourrait inclure l'établissement de directives sur le renforcement des capacités des enseignants en matière d'évaluation en classe des acquis des élèves et le renforcement des capacités techniques des services ministériels. Un

⁷ Banerji, Rukmini et Michael Walton. « *What Helps Children to Learn? Evaluation of Pratham's Read India Program in Bihar & Uttarakhand* ». New Delhi: Pratham, 2011. Save the Children. « *Beyond School Walls: A Boost for Readers* ». Fairfield CT: Save the Children, 2013.

⁸ Care, Esther, Alvin Vista et Helyn Kim. « *Optimizing Assessment for All* ». Washington, DC: Brookings Institution, 2018.

examen des bonnes pratiques de mise en œuvre de ces deux outils pourrait aussi s'avérer utile en l'occurrence. Des innovations peu coûteuses sur le renforcement durable des compétences des services ministériels et des enseignants pourraient aussi s'avérer utiles. Ces efforts pourraient inclure la création de biens publics mondiaux favorisant la transposition des acquis d'un pays à l'autre.

Biens mondiaux à l'appui des systèmes d'évaluation de l'apprentissage dans les pays les plus marginalisés, y compris les PFC et les élèves les plus marginalisés qui y habitent. Les activités pourraient inclure l'aide aux PFC qui ne réalisent pas d'évaluations des résultats scolaires — afin qu'ils puissent participer aux évaluations transnationales répondant aux normes de qualité du PME — et des directives sur l'harmonisation des priorités des donateurs concernant les évaluations, y compris un examen de la rentabilité des évaluations de l'apprentissage, et un examen des succès remportés par les PFC dans leur participation aux évaluations nationales et internationales. Les investissements dans ce domaine pourraient aussi renforcer les mécanismes conçus pour aider les pays à utiliser les données existantes aux fins de la communication des résultats à l'échelle mondiale et de l'élaboration de solutions technologiques peu coûteuses pour la collecte des données dans les PFC et auprès des élèves handicapés ou déplacés en raison des conflits.

Bien mondiaux à l'appui du recours systémique aux données d'évaluation de l'apprentissage. Cet enjeu recoupe dans une certaine mesure celui des travaux thématiques sur les données du KIX, mais les PDP et les experts consultés se sont dits convaincus qu'il devrait aussi être pris en compte. Les échanges régionaux et internationaux sur les bonnes pratiques de présentation et d'utilisation des données d'évaluation de l'apprentissage pourraient s'avérer utiles aux PDP. IL serait aussi utile de procéder à un examen de la façon dont les données sur cette question (évaluations en classe, examens et évaluations à grande échelle) influent sur les politiques des pays à revenu faible ou moyen. Une approche systématique « tester-intervenir-tester à nouveau » basée sur les évaluations de l'apprentissage pourrait offrir un nouvel éclairage sur la façon dont les pays peuvent intégrer ces évaluations dans leurs programmes d'éducation. La mise en place de nouveaux moyens de communication et de visualisation des données pourrait améliorer l'utilisation des données sur l'évaluation de l'apprentissage.

Pour veiller à ce que les investissements consacrés au mécanisme KIX dans les domaines précités répondent aux besoins et au contexte, plusieurs types d'intervention sont requis, y compris :

- renforcement des capacités par le transfert des connaissances et l'échange d'expériences sur les enjeux sur lesquels il existe des données probantes suffisantes ;
- collecte de données probantes et évaluation de ce qui fonctionne sur les questions pour lesquelles il existe des exemples de solutions, mais sur lesquelles nous devons effectuer une synthèse plus approfondie afin de bâtir une solide assise factuelle ;
- innovations relatives à l'évaluation de l'apprentissage dans des domaines pour lesquels on constate un besoin de réflexion et de solutions nouvelles.

1. Introduction

L'utilisation efficace des données d'évaluation des acquis de l'apprentissage revêt une importance fondamentale pour les systèmes d'éducation. À défaut de disposer de données sur le processus d'apprentissage des élèves, il est en effet impossible de savoir si le système éducatif offre une éducation de qualité pour tous les enfants. L'évaluation consiste à « recueillir et à évaluer des données qui nous

renseignement sur ce que les élèves connaissent et comprennent, et sur ce qu'ils peuvent faire pour prendre une décision éclairée sur les prochaines étapes de leur cheminement éducatif »⁹. L'évaluation de l'apprentissage peut être un outil efficace pour améliorer les systèmes d'éducation nationale lorsqu'elle s'appuie sur des données contextuelles. Un système d'évaluation de l'apprentissage englobe l'ensemble des politiques, institutions, outils et pratiques d'évaluation nécessaires à la collecte et à l'utilisation des informations sur les performances des élèves dans le système d'éducation¹⁰. Les composantes d'un tel système varient selon le pays, et peuvent inclure des évaluations conçues et réalisées par les enseignants pour étayer leur pratique (c'est-à-dire, des évaluations réalisées en classe), des examens nationaux administrés à tous les élèves, et des évaluations basées sur des échantillons, utilisées principalement aux fins de la prise de décisions (évaluations à grande échelle). Un système rigoureux d'évaluation de l'apprentissage peut aider les élèves, leurs parents, les enseignants et les administrateurs d'établissements scolaires à suivre les progrès réalisés et à intervenir pour aider ceux qui ont des difficultés d'apprentissage.

L'évaluation de l'apprentissage est au cœur du programme mondial d'éducation défini dans les objectifs de développement durable (ODD) et par le cadre d'action Éducation 2030. Le quatrième objectif de développement durable¹¹ énonce un ensemble ambitieux de cibles, et notamment l'amélioration des résultats d'apprentissage pour tous, de la petite enfance à l'âge adulte. La moitié des dix objectifs de l'ODD 4 liés à l'éducation ont trait à l'apprentissage et aux compétences¹². Cette insistance des ODD sur l'apprentissage est le résultat d'une vaste campagne menée par les pouvoirs publics, la société civile, les organisations d'enseignants et les bailleurs, y compris le Partenariat mondial pour l'éducation (PME), en faveur d'une accélération des progrès dans le secteur de l'éducation, et de l'engagement de tous ces acteurs à obtenir de meilleures données sur l'apprentissage¹³.

Nous sommes désormais investis d'un mandat à caractère mondial en faveur d'un renforcement des systèmes d'évaluation de l'apprentissage conforme aux ODD. En utilisant les données existantes, nous pouvons estimer le nombre d'enfants qui n'arrivent pas à satisfaire aux normes d'apprentissage minimales en calcul et en lecture (6 sur 10, soit 617 millions d'enfants dans le monde)¹⁴. Cependant, ces chiffres restent très approximatifs puisque plusieurs pays à revenu faible ou moyen, y compris les pays en développement partenaires du PME (PDP) ne disposent pas de données nationales récentes sur les résultats d'apprentissage. Les pays africains, les pays les moins avancés, les pays en développement enclavés, les petits États insulaires en développement et les pays en proie à un conflit affichent les taux les plus bas de participation aux évaluations à grande échelle¹⁵. Bien que les estimations mondiales puissent attirer l'attention sur le problème, elles doivent pouvoir s'appuyer sur des systèmes nationaux,

⁹ Clarke, 2012, 1.

¹⁰ Clarke, 2012.

¹¹ L'ODD 4 consiste à « assurer l'accès de tous à une éducation de qualité, sur un pied d'égalité, et promouvoir les possibilités d'apprentissage tout au long de la vie ».

¹² Les objectifs 4.1, 4.2, 4.4, 4.6 et 4.7 et les indicateurs correspondants sont liés aux résultats d'apprentissage.

¹³ Groupe de personnalités de haut niveau chargé d'étudier le programme de développement pour l'après-2015. « *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development* ». New York: Nations Unies, 2013 ; Comité de réflexion sur la métrique des apprentissages. « *Toward Universal Learning: Recommendations from the Learning Metrics Task Force* ». Montréal et Washington DC : ISU et Brookings Institution. 2013 ; Groupe thématique 4 du SDSN. « *The Future of Our Children: Lifelong, Multi-Generational Learning for Sustainable Development* ». New York: SDSN, 2014.

¹⁴ Institut de statistique de l'UNESCO. « *6 Out of 10 Children and Adolescents Are Not Learning a Minimum in Reading and Math* ». Montréal: ISU, UNESCO, 2017a.

¹⁵ Institut de statistique de l'UNESCO. « *Investment Case for Expanding Coverage and Comparability for Global Indicator 4.1.1* ». Montréal: ISU, UNESCO, 2017b.

régionaux et internationaux d'évaluation de l'apprentissage à la fois solides et adaptés, utiles aux administrations publiques chargées de leur gestion.

Le PME et beaucoup d'autres organismes de développement ont déployé des efforts pour renforcer les systèmes d'évaluation de l'apprentissage, notamment en aidant les pays à améliorer leurs systèmes et à produire des biens publics mondiaux qui transcendent les frontières nationales et qui peuvent servir à tous les pays en développement, qu'ils soient ou non membre du PME¹⁶. Les biens mondiaux sont les institutions, les mécanismes et les résultats qui procurent des avantages quasi universels, et transcendent les frontières et les générations¹⁷, c'est-à-dire les outils, produits et méthodes — y compris les données, les outils d'évaluation, les normes et les résultats de recherches — qui, une fois élaborés à l'issue d'une intervention particulière, peuvent être adaptés pour créer un outil ou une approche applicable à d'autres contextes, à condition d'y apporter les adaptations requises¹⁸. Le présent document décrit les efforts déployés pour renforcer les systèmes d'évaluation de l'apprentissage, déterminer les lacunes qui persistent en matière de connaissances et d'innovation, et proposer un ensemble de domaines où le PME pourrait investir par l'intermédiaire de son nouveau mécanisme de partage de connaissances et d'innovations (KIX) pour renforcer ces systèmes.

2. Préparation du document et processus de consultation

Le PME a chargé un auteur principal d'examiner ses documents et ses plans sectoriels portant sur le renforcement des systèmes d'évaluation de l'apprentissage, de réaliser une étude documentaire, de consulter les PDP et les experts, et de préparer le présent rapport. Collaborant avec les spécialistes seniors de l'éducation du PME chargés des systèmes d'évaluation de l'apprentissage et avec d'autres membres du Secrétariat du PME, cet auteur a préparé en mai 2018 un document initial de discussion qui résumait les obstacles à la mise en place de systèmes solides d'évaluation de l'apprentissage, les biens mondiaux disponibles, les carences de ces biens mondiaux mises en lumière par l'étude documentaire, et les moyens par lesquels le PME pourrait intervenir par le biais du mécanisme KIX. Ce document a été examiné par le Secrétariat du PME. Une version révisée a été traduite en français, et les deux versions ont été transmises aux représentants des PDP et à des experts internationaux représentant les universités, la société civile, des fondations, des donateurs et des institutions multilatérales.

Des consultations ont été organisées au cours d'un atelier régional du Réseau des éducateurs pour la transformation (TALENT) de l'UNESCO Dakar tenu à Dar es Salaam, en juillet 2018, et d'un autre tenu dans la région Asie-Pacifique par le Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique (NEQMAP) de l'UNESCO Bangkok, à Pénang, en septembre. Ces ateliers rassemblaient quelque 65 représentants officiels de 30 PDP, chargés typiquement de la direction des évaluations de l'apprentissage, des programmes d'enseignement et de la formation des enseignants.

Le Secrétariat du PME a reçu des retours d'informations en présentiel de 30 PDP, et des observations écrites de 7 PDP et de 24 experts internationaux. Les résultats de ces consultations sont présentés avec

¹⁶ Voir à la section 4 la liste des nombreux biens mondiaux existants qui servent à renforcer les systèmes d'évaluation de l'apprentissage.

¹⁷ UNESCO. Document d'orientation n° 34 : « *Fulfilling our collective responsibility: Financing global public goods in education* ». Paris: UNESCO, 2018.

¹⁸ Commission internationale pour le financement de possibilités d'éducation dans le monde. « *La génération d'apprenants : Investir dans l'éducation pour un monde en pleine évolution* ». Washington DC: Commission pour l'éducation, 2016.

les résultats de recherche. Pour les enjeux faisant l'objet d'un consensus de l'ensemble des intervenants ou pour lesquels les retours d'information permettaient de clarifier le contenu du rapport initial, les retours d'information ont été intégrés directement dans le présent document. Pour ceux faisant l'objet de points de vue divergents, ces points de vue ont été débattus et exposés dans le présent document, accompagnés, le cas échéant, de recommandations sur les suites à donner.

3. Obstacles à la mise en place de systèmes solides d'évaluation de l'apprentissage

Plusieurs raisons permettent d'expliquer l'absence de systèmes solides d'évaluation de l'apprentissage, en particulier dans les pays les plus vulnérables. Les travaux et les consultations menées auprès des PDP et des experts internationaux ont mis en lumière quatre obstacles principaux pouvant expliquer cette situation.

La première difficulté découle de la piètre qualité des outils d'évaluation et d'un manque de compétences techniques dans les domaines de la conception et de l'administration des évaluations, et de l'analyse de leurs résultats. Elle présente un obstacle majeur. Il importe de disposer de bons outils de mesure de la qualité (tests). Pour évaluer la qualité d'un outil de mesure, il faut notamment en vérifier la précision ou la fiabilité, ainsi que la validité ou la valeur des données qu'il permet de recueillir. Pour les outils utilisés dans différents contextes culturels, la validité interculturelle constitue aussi un facteur important.

L'adaptation au contexte culturel constitue un autre aspect important de la qualité des outils. Par exemple, les évaluations écrites administrées à l'école peuvent constituer un moyen approprié de mesurer les compétences en lecture dans les pays où les enfants fréquentent l'école et sont capables de lire. Cependant, dans les pays où ces conditions ne sont pas réunies, le recours à des évaluations orales pourrait s'avérer plus approprié. La collecte de données par le biais d'enquêtes menées auprès des ménages peut servir à compléter les informations recueillies par les établissements scolaires. Cependant, les PDP et les experts internationaux ont signalé qu'il arrive souvent que les pays décident de la méthode d'évaluation à retenir avant d'engager un dialogue sur ces types d'enjeux.

Des experts hautement qualifiés sont nécessaires pour la conception, l'administration et l'analyse des données des évaluations à grande échelle de l'apprentissage et des examens pour veiller à ce qu'ils répondent aux normes de qualité. Or, ces experts risquent de ne pas être disponibles dans le pays ou dans la région, et il devient alors nécessaire de recourir à des experts internationaux dont le coût sera beaucoup plus élevé que celui du personnel local. Les parties prenantes consultées dans le cadre de la préparation du présent document ont fait valoir qu'il est parfois difficile d'obtenir l'ensemble des outils ou des méthodes nécessaires aux évaluations à grande échelle, et qu'il est en conséquence difficile de compléter les travaux déjà réalisés. Compte tenu de ces exigences, les évaluations de l'apprentissage peuvent s'avérer coûteuses. On estime que les évaluations à grande échelle coûtent 500 000 dollars par pays en moyenne pour la collecte des données et l'assistance technique¹⁹. Le coût d'une évaluation de

¹⁹ ISU, 2017a.

l'apprentissage précoce de la lecture peut varier grandement, oscillant entre 150 000 et 300 000 dollars²⁰.

Les parties prenantes consultées dans le cadre de la présente étude ont indiqué qu'il arrive souvent que les activités de renforcement des capacités et de formation des décideurs, des chercheurs ou des enseignants portant sur un type ou l'autre d'évaluation soient de nature informelle, de courte durée et de qualité variable. Par ailleurs, il peut parfois arriver qu'elles ne soient pas reconnues officiellement ou agréées.

Les évaluations en classe sont souvent considérées comme peu prioritaires en raison de la piètre qualité des outils disponibles et du manque de formation des enseignants à la conception et à l'utilisation de ces outils. Il arrive aussi que les écoles et les enseignants ne disposent pas des ressources et des options nécessaires pour procéder aux améliorations requises à partir des informations fournies par les évaluations en classe. Ainsi, les enseignants, les élèves et leurs familles ne savent pas si l'apprentissage progresse comme prévu. Plusieurs PDP ont indiqué que dans les classes surpeuplées, pouvant compter de 60 à 80 élèves, il est difficile pour les enseignants d'apprendre à connaître leurs élèves suffisamment pour évaluer d'une manière cohérente la qualité de leur apprentissage.

Le deuxième obstacle est posé par le manque de cohésion entre les divers types d'évaluation de l'apprentissage utilisés, et par les difficultés que présente l'intégration de ces évaluations aux systèmes éducatifs nationaux. Cet obstacle découle du fait que ce sont les taux de scolarisation, plutôt que les acquis scolaires, qui retiennent avant tout l'attention des systèmes d'éducation. Ainsi, l'incitation à l'investissement dans les outils de mesure et d'amélioration des résultats scolaires reste souvent faible²¹. Le décalage entre les programmes, la pédagogie et l'évaluation est aussi important, les programmes s'avérant souvent trop avancés pour la plupart des élèves, la formation des enseignants étant basée davantage sur la théorie que sur la pratique, et les politiques d'évaluation étant exclusivement centrées sur deux ou trois domaines²².

On observe aussi parfois un décalage au sein du système d'évaluation de l'apprentissage entre les multiples évaluations réalisées dans un pays donné. Un des experts des PDP consultés a mentionné que les enseignants utilisant les évaluations en classe se montraient surpris de constater que leurs élèves obtenaient des résultats nettement inférieurs à ceux prévus lors des examens. La recherche d'un compromis entre la nécessité d'évaluer et la poursuite des objectifs fondamentaux d'un système éducatif peut aussi représenter un défi. Des évaluations trop fréquentes risquent d'accaparer une part précieuse du temps de classe, mais des évaluations trop rares peuvent priver les enseignants et les décideurs des informations dont ils ont besoin pour s'acquitter de leurs tâches.

En insistant trop fortement sur des examens à forts enjeux, qui mettent souvent l'accent sur la mémorisation et se concentrent sur quelques domaines particuliers, on empêche les enseignants de couvrir l'ensemble du programme scolaire. Dans les pays où le nombre de places dans le secondaire est

²⁰ Wagner, Daniel, Martin Lockheed, Ina Mullis, Michael O. Martin, Anil Kanjee, Amber Gove et Amy Dowd. « The Debate on Learning Assessments in Developing Countries ». *Journal of Comparative and International Education*, 42(3): 509-545, 2012.

²¹ Pritchett, Lant. « *Creating Education Systems Coherent for Learning Outcomes: Making the Transition* ».

²² Institut international de l'UNESCO pour le renforcement des capacités en Afrique. « *Teaching policies and learning outcomes in Sub-Saharan Africa: Issues and Options* ». Addis-Abeba: Institut international de l'UNESCO pour le renforcement des capacités en Afrique, 2016 ; Anderson, Kate, Seamus Hegarty, Martin Henry, Helyn Kim et Esther Care. « *Breadth of Learning Opportunities: A Fresh Approach to Evaluating Education Systems* ». Washington DC: Brookings Institution, 2018.

limité, cette pratique peut devenir un outil d'exclusion plutôt qu'un outil visant à améliorer l'apprentissage. Les données des examens peuvent devenir une source accessible et utile d'informations sur l'ensemble des élèves. Or, elles sont rarement utilisées pour éclairer les politiques et les décisions en classe.

Les nombreuses sources de financement des systèmes d'éducation, y compris les organisations nationales et les bailleurs internationaux, ont des priorités différentes. Ce manque d'alignement risque de nuire à l'efficacité des systèmes d'évaluation de l'apprentissage. Le financement des évaluations peut se montrer aléatoire, et favoriser ainsi une approche au cas par cas où les évaluations ne sont réalisées que lorsque les financements le permettent. Ce point de vue n'était cependant pas partagé par l'ensemble des PDP et des experts internationaux. Certains ont jugé que les budgets alloués suffisaient pour couvrir le coût des évaluations, et que ces coûts ne représentaient de toute manière qu'une mince part du budget total de l'éducation. Le manque de communication entre les divers services publics chargés de prendre les décisions dans les ministères de l'éducation a aussi été reconnu comme un défi.

Le troisième obstacle découle du fait que les évaluations de l'apprentissage ne prennent souvent pas en compte les enfants les plus marginalisés, y compris ceux qui sont handicapés ou qui ne sont pas scolarisés. Même lorsque les pays disposent de données nationales sur le degré global d'apprentissage des enfants, ces données risquent malgré tout ne pas suffire pour analyser le processus d'apprentissage de différents groupes d'élèves — garçons et filles, riches et pauvres, enfants souffrant d'un handicap, enfants déplacés par un conflit, ou enfants qui ne parlent pas la langue dominante à la maison. Pour être utiles, les données des évaluations à grande échelle et des examens publics devraient être scrutées à l'aune de facteurs contextuels afin de déterminer le processus d'apprentissage de différents groupes d'élèves. Or, il arrive souvent que les données collectées ne soient pas analysées suffisamment en profondeur pour fournir des indications utiles sur les moyens d'intervention.

On néglige souvent de soumettre les enfants handicapés à des évaluations appropriées en classe, et les ressources requises pour modifier comme il se doit l'environnement et la pédagogie sont souvent insuffisantes. Comme on l'a vu plus haut, les classes surpeuplées, qui tendent à être plus fréquentes dans les collectivités marginalisées, peuvent empêcher les enseignants d'évaluer correctement les élèves et d'utiliser les données pour apporter les améliorations requises²³. L'absentéisme élevé des élèves comme des enseignants peut aussi empêcher les enseignants de recueillir régulièrement des données d'évaluation en classe.

Le quatrième obstacle découle d'une utilisation insuffisante des résultats des évaluations de l'apprentissage. Ce problème est dû à la manière dont les évaluations sont conçues et communiquées, et aux compétences et aux ressources disponibles pour procéder aux changements nécessaires. Pour être valable, une évaluation doit principalement être adaptée au but recherché, c'est-à-dire qu'elle doit fournir les informations nécessaires à ceux qui en feront usage²⁴. L'évaluation peut poursuivre plusieurs buts, y compris l'aide à l'apprentissage en classe, l'élaboration de politiques à l'échelle du système, la reddition de comptes, l'évaluation de projets spéciaux ou de projets pilotes, la certification, la promotion et les transferts. Idéalement, ces buts sont définis avant le début d'une évaluation, et les systèmes d'éducation consacrent des ressources suffisantes aux diverses composantes d'un système d'évaluation de l'apprentissage. En pratique cependant, on met d'ordinaire l'accent sur les évaluations

²³ UNESCO. « Rapport mondial de suivi sur l'EPT : Atteindre les marginalisés ». Paris: UNESCO, 2010.

²⁴ Archer, Elizabeth. « The Assessment Purpose Triangle: Balancing the Purposes of Educational Assessment ». *Frontiers in Education*. 2 (41), 2017.

dont les enjeux sont les plus grands, c'est-à-dire, en général, les examens de fin de cycle²⁵. Certaines composantes d'un système d'évaluation de l'apprentissage peuvent aussi remplir plus d'une fonction, permettant ainsi de réaliser des économies. Cependant, selon les parties prenantes consultées dans le cadre de la présente étude, on néglige d'ordinaire de tirer parti de cette possibilité.

Le fait que les données sur les résultats d'apprentissage ne soient souvent pas utilisées d'une manière qui conduit à l'amélioration de l'apprentissage est une préoccupation qu'ont partagée la presque totalité des PDP et des experts internationaux consultés. Ces derniers ont fait état d'un décalage (de l'absence de boucles de rétroaction) entre les données et la pratique. Les évaluations en classe ont aussi suscité des inquiétudes semblables puisque les ressources et les compétences requises pour procéder aux changements nécessaires sont limitées. Le même problème peut aussi se poser au niveau du système tout entier. Il peut enfin arriver qu'on manque de personnel pour analyser les données, ou que les données soient effectivement analysées, mais qu'elles ne se retrouvent pas entre les mains de ceux qui peuvent effectuer les changements nécessaires. Il arrive aussi que les ministères de l'Éducation recueillent d'abord les données, et décident ensuite de la manière de les utiliser, au lieu de définir d'entrée de jeu les questions auxquelles ils souhaitent apporter des réponses.

Enfin, les données sur l'apprentissage sont souvent analysées sous l'angle des notes moyennes, l'objectif étant de hausser les moyennes nationales, ou de faire passer plus d'élèves dans la catégorie « compétent », ce qui conduit à exclure les élèves en difficulté et les apprenants plus avancés²⁶.

4. Appui du PME aux systèmes d'évaluation de l'apprentissage

En raison de tous ces obstacles, le PME a fait de la mise en place de systèmes solides d'évaluation de l'apprentissage une priorité, comme en font foi son cadre de résultats et son modèle de financement. Améliorer les résultats d'apprentissage et les rendre plus équitables est le premier des trois buts énoncés dans le Plan stratégique 2020 du PME²⁷, et le cadre de résultats et le modèle de financement du PME nécessitent ainsi un suivi des résultats d'apprentissage. En 2015, le PME a élaboré son premier cadre de résultats pour la période 2016-2020²⁸. Ce cadre comporte un indicateur relatif aux acquis scolaires et deux autres relatifs aux systèmes d'évaluation de l'apprentissage :

Indicateur 1 : Proportion des PDP enregistrant une amélioration des acquis scolaires (éducation de base). Cet indicateur utilise les données d'évaluations des acquis scolaires représentatives au plan national pour déterminer le nombre de pays qui ont connu une amélioration statistiquement significative des résultats scolaires au primaire et au premier cycle du secondaire.

²⁵ Wagner *et al.*, 2012 ; Comité de réflexion sur la métrique des apprentissages. « *Toward Universal Learning: Implementing Assessment to Improve Learning* ». Montréal et Washington DC: ISU et Brookings Institution, 2014.

²⁶ Wagner, Daniel, Sharon Wolf et Robert F. Boruch, eds. « *Learning at the Bottom of the Pyramid: Science, Measurement and Policy in Low-Income Countries* ». Paris: UNESCO IIEP, 2018.

²⁷ Partenariat mondial pour l'éducation. 2016. « Plan stratégique 2020 du PME ». Partenariat mondial pour l'éducation 2020. Washington DC: PME, 2016.

²⁸ Partenariat mondial pour l'éducation. « Cadre de résultats du PME pour 2016-2020 ». Partenariat mondial pour l'éducation. Washington DC: PME, 2017.

Indicateur 15 : proportion de PDP dont le système d'évaluation des acquis scolaires dans l'éducation de base répond à des normes de qualité. Les normes de qualité concernent notamment le contexte dans lequel est réalisée une évaluation, la qualité des outils et des processus d'évaluation et la mesure dans laquelle l'évaluation s'aligne sur le programme scolaire national.

Indicateur 20 : proportion des financements appuyant les systèmes d'information sur la gestion de l'éducation (SIGE) ou les systèmes d'évaluation des apprentissages. Cet indicateur s'intéresse au nombre de financements pour la mise en œuvre de programmes sectoriels de l'éducation (ESPIG) qui appuient les systèmes d'évaluation de l'apprentissage ou les systèmes d'information sur la gestion de l'éducation.

On trouvera à l'annexe A de plus amples informations sur ces indicateurs et sur les progrès réalisés à ce jour par les PDP.

Il est possible en utilisant les données recueillies pour l'Indicateur 20 de déterminer le nombre de PDP bénéficiant d'une aide du PME pour le renforcement de leurs systèmes d'évaluation des acquis scolaires. Au 30 juin 2017, 29 (ou 71 %) des 41 ESPIG actifs investissaient dans des activités d'évaluation de l'apprentissage²⁹. Les activités bénéficiant de ces financements comprennent l'élaboration et la mise en œuvre de systèmes d'évaluation à l'échelle de la classe et d'évaluation nationale au Bangladesh, et la mise sur pied d'une agence indépendante chargée des évaluations nationales en République démocratique du Congo.

Le modèle de financement du PME (adopté en 2014 par le Conseil d'administration pour la période de reconstitution 2015-2018) exige des pays qui présentent une requête de financement pour la mise en œuvre de leur programme d'éducation qu'ils disposent d'un système de suivi des acquis scolaires ou d'un plan pour la mise en place d'un tel système³⁰. En 2017, le Conseil du PME a décidé que si les financements disponibles ne suffisaient pas pour mettre en place ces systèmes, on utiliserait les fonds du PME pour le faire. Le nouveau modèle de financement permet aussi au PME de fournir des financements basés sur les résultats, encourageant ainsi les pays à établir et à atteindre leurs propres cibles d'apprentissage. Pour recevoir la première tranche de 70 % des financements du PME, chaque PDP doit remplir plusieurs conditions importantes, comme expliqué plus haut. Le décaissement de la dernière tranche de 30 % dépend de données concrètes démontrant les progrès réalisés dans le secteur de l'éducation, y compris les gains enregistrés dans le domaine de l'apprentissage. Les autorités publiques, en consultation avec leurs partenaires au sein du Groupe local des partenaires de l'éducation (GLPE), doivent définir une stratégie de transformation visant à améliorer les résultats scolaires et décrivant les mesures qui seront prises pour lever les obstacles qui nuisent à l'apprentissage dans leurs contextes particuliers. Par exemple, la République démocratique du Congo a lié les financements à l'amélioration du niveau de lecture des élèves du cycle primaire. L'engagement à améliorer les acquis scolaires constitue donc un des moyens par lesquels les PDP peuvent obtenir l'accès au financement axé sur les résultats du Partenariat.

²⁹ Ce chiffre n'inclut pas les fonds mis en commun à l'échelon sectoriel.

³⁰ Les examens publics et la délivrance de diplômes n'entrent pas en ligne de compte dans ce cas.

Le PME a aussi appuyé des activités de portée mondiale et régionale visant à renforcer l'évaluation de l'apprentissage. Le Programme d'activités mondiales et régionales (AMR), qui a été officiellement conclu, comportait deux financements axés sur les systèmes d'évaluation de l'apprentissage qui totalisaient près de 1,5 million de dollars. Le premier de ces financements appuyait un projet de l'ISU réalisé entre 2013 et 2015 et qui avait pour buts d'élaborer des méthodes permettant de comparer les évaluations de lecture entre les régions, de définir les meilleures pratiques en matière d'évaluation des compétences en lecture dans les petites classes et de mettre en place un inventaire mondial des évaluations d'apprentissage. Le deuxième financement appuyait les activités du Réseau de suivi de la qualité de l'éducation dans la région Asie-Pacifique (NEQMAP) pour recueillir des données probantes sur le plan régional et renforcer les capacités dans ce domaine. Ces activités ont été mises en œuvre par l'UNESCO Bangkok de 2014 à 2016.

L'initiative Évaluation au service des apprentissages (A4L) a été lancée en juillet 2017, à la suite de ces activités, pour renforcer la capacité des systèmes nationaux d'évaluation de l'apprentissage, dans le cadre d'un projet pilote de mise en œuvre de l'approche d'échange de connaissances et de bonnes pratiques. L'Initiative A4L dispose d'un budget d'environ 3 millions de dollars et comporte trois volets : 1) outils de diagnostic des systèmes d'évaluation de l'apprentissage, à publier à la suite de tests pilotes effectués dans trois PDP en 2018 et 2019 ; 2) aide au NEQMAP et au réseau TALENT (Enseignement et apprentissage : Réseau des éducateurs pour la transformation) en Afrique Subsaharienne pour le renforcement des capacités, la réalisation d'analyses et le partage des connaissances ; 3) tour d'horizon sur la mesure des compétences du XXI^e siècle et les outils nécessaires à cette fin (en collaboration avec le Centre pour l'éducation universelle de la Brookings Institution).

Le PME appuie vigoureusement les systèmes d'évaluation de l'apprentissage en les intégrant à son Plan stratégique 2020, en aidant les PDP à assurer le suivi des acquis scolaires, en participant aux efforts déployés à l'échelle mondiale et régionale pour améliorer les évaluations de l'apprentissage, et en renforçant les capacités des systèmes nationaux d'évaluation de l'apprentissage par le biais de l'Initiative A4L. L'Initiative A4L a en particulier planté le décor pour le mécanisme KIX en favorisant une approche systémique du renforcement des systèmes d'évaluation de l'apprentissage. Le PME s'emploie à mettre en place un outil qui aidera les pays à analyser leurs systèmes d'évaluation de l'apprentissage et servira ainsi de complément aux efforts de renforcement de ces systèmes. Le Partenariat prête aussi assistance aux réseaux régionaux de renforcement des compétences et aide ainsi les pays à se doter de systèmes d'évaluation répondant aux besoins des programmes nationaux et des priorités stratégiques. Les financements pour le mécanisme KIX prendront le relai de ces efforts en favorisant le partage des connaissances, en appuyant les nouveaux biens mondiaux et en tirant parti de ceux qui existent, tel que décrit dans la section suivante.

5. Biens mondiaux et innovations relatifs à l'évaluation de l'apprentissage

Les biens mondiaux relatifs à l'évaluation de l'apprentissage dans les pays en développement se répartissent en trois grandes catégories : évaluations à grande échelle de l'apprentissage ; outils et initiatives de renforcement des capacités ; innovations relatives à l'évaluation de l'apprentissage.

Évaluations à grande échelle de l'apprentissage

Les biens mondiaux de cette catégorie englobent les outils et les études, les données, les réseaux et les mécanismes de partage des connaissances.

Outils et études

Les évaluations à grande échelle servent au suivi et à la production d'informations aux décideurs et aux praticiens sur les niveaux globaux de performance, sur l'évolution de ces niveaux au fil du temps, et sur les liens qui existent entre diverses caractéristiques des élèves, de leurs familles, de leurs écoles et de leurs collectivités et les niveaux de performance. Ces évaluations peuvent poursuivre de multiples objectifs, mais elles servent d'ordinaire à fournir des informations sur la santé générale des systèmes et à décider des politiques à mettre en œuvre et de l'affectation des ressources.

Les évaluations à grande échelle peuvent être nationales ou internationales. Les évaluations nationales sont propres à chaque pays et s'alignent d'ordinaire sur le programme scolaire national. Le mécanisme de suivi des acquis scolaires (MLA) de l'UNESCO et de l'UNICEF a été l'un des premiers biens mondiaux venant à l'appui des systèmes nationaux d'évaluation de l'apprentissage dans les pays à revenu faible ou moyen. Dans la foulée de la Conférence mondiale sur l'éducation pour tous 1990 organisée à Jomtien (Thaïlande), l'UNESCO et l'UNICEF ont mené une opération pilote visant à mettre en place des systèmes nationaux d'évaluation en collaboration avec la Chine, la Jordanie, le Mali, Maurice et le Maroc au milieu des années 1990³¹. Les lignes directrices et l'approche qui en ont découlé ont servi dans toute l'Afrique subsaharienne et en Asie, et plusieurs PDP continuent d'utiliser des versions révisées du MLA en guise d'outil national d'évaluation³². Les évaluations internationales autorisent les comparaisons entre les pays participants. Elles sont souvent le produit d'un processus consensuel et tiennent compte des programmes scolaires nationaux de l'ensemble des pays participants. Les deux types d'évaluations peuvent aussi permettre d'établir des comparaisons dans le temps, lorsqu'elles sont conçues à cette fin.

Un pays qui souhaite mesurer l'apprentissage en adaptant un outil existant ou en participant à une étude avec d'autres pays peut choisir entre plusieurs types d'outils qui peuvent être considérés comme des biens globaux, comme le montre le tableau 1. Ces outils comprennent les études comparatives internationales et régionales. Les méthodes d'évaluation mises à disposition aux fins d'adaptation comprennent les évaluations de la lecture et du calcul dans les petites classes (EGRA et EGMA) et les évaluations citoyennes. Ces évaluations ne fournissent pas de données comparables d'un pays à l'autre, mais elles favorisent des économies d'échelle pour les pays qui souhaitent évaluer l'apprentissage sans avoir à bâtir un système à partir de zéro. Le module sur les compétences fondamentales d'apprentissage des enquêtes par grappes à indicateurs multiples (MICS) de l'UNICEF est un autre outil pertinent ; il s'agit d'une nouvelle méthode d'enquête auprès des ménages conçue pour fournir des données relatives à l'ODD 4.1.1(a). Les résultats sont comparables à l'échelle mondiale et seront disponibles pour un nombre élevé de PDP à la fin de 2018. Cette évaluation prend en compte les enfants non scolarisés et permet de réaliser une analyse détaillée de l'équité fondée sur les informations sociodémographiques substantielles recueillies auprès des ménages, ouvrant ainsi la voie à l'analyse intersectorielle en raison des informations recueillies portant sur la nutrition, la santé, le travail des enfants et les châtements corporels.

³¹ Chinapah, Vinayagum. « *Monitoring Learning Achievement: Toward Capacity Building* ». Paris: UNESCO, 1997.

³² Friedman, Tim, Ursula Schwantner, Jeaniene Spink, Naoko Tabata et Charlotte Waters. « *Improving Quality Education and Children's Learning Outcomes and Effective Practices in the Eastern and Southern Africa Region* ». Nairobi: UNICEF/ Bureau régional de l'Afrique de l'Est et de l'Afrique australe, 2016.

Tableau 1 : Outils plurinationaux d'évaluation de l'apprentissage

	Nom	Sigle	Niveau ou âge des élèves évalués	Matières ou domaines évalués	Organisme chargé de l'administration
Les résultats peuvent être comparés d'un pays à l'autre	Programme international de recherche en lecture scolaire	PIRLS	Niveau 4	Lecture	Association internationale pour l'évaluation du rendement scolaire (IEA)
	Étude internationale sur les tendances de l'enseignement des sciences et des mathématiques	TIMSS	Niveau 4 Niveau 8	Mathématiques Sciences	IEA
	Étude internationale sur l'éducation à la citoyenneté	ICCS	Niveau 8	Éducation civique et citoyenne, citoyenneté mondiale	IEA
	Programme international pour le suivi des acquis des élèves	PISA	15 ans	Lecture Mathématiques Sciences Domaines optionnels (par exemple, résolution de problèmes en collaboration, en 2015)	Organisation de coopération et de développement économiques
	Laboratoire latino-américain pour l'évaluation de la qualité de l'éducation	LLECE	Niveau 3 Niveau 6	Lecture et écriture Mathématiques Sciences naturelles (niveau 6)	Bureau régional de l'UNESCO pour l'éducation en Amérique latine et dans les Caraïbes
	Enquête par grappes à indicateurs multiples : module sur les compétences fondamentales d'apprentissage	MICS	7 à 14 ans	Lecture Calcul	UNICEF
	Programme d'analyse des systèmes éducatifs de la CONFEMEN	PASEC	Niveau 2 Niveau 6	Lecture Mathématiques	CONFEMEN
	Évaluation	PILNA	Niveau 4	Lecture	Programme

	régionale des aptitudes à la lecture, à l'écriture et au calcul de la communauté du Pacifique		Niveau 6	Calcul	d'évaluation de la qualité de l'éducation
	Consortium d'Afrique australe et orientale pour le suivi de la qualité de l'éducation	SEACMEQ	Niveau 6	Lecture Mathématiques Connaissances sur le VIH/SIDA	SEACMEQ
	Évaluation des aptitudes à la lecture, à l'écriture et au calcul	LaNA	Niveau 4 (ou 5 ou 6, le cas échéant)	Compétences de base en lecture Compétences de base en mathématiques	IEA
	Initiative PISA pour le développement*	PISA-D	15 ans (incluant les enfants non scolarisés dans la plupart des pays)	Lecture Mathématiques Sciences	PISA
	Initiative pour la mesure de l'apprentissage à l'école primaire en Asie du Sud-est*	SEA-PLM	Niveau 5	Lecture et écriture Mathématiques Citoyenneté mondiale	Association des Ministres de l'Éducation d'Asie du Sud-est / Bureau régional de l'UNICEF pour l'Asie de l'Est et le Pacifique
Les résultats ne se comparent pas d'un pays à l'autre	Évaluations citoyennes (par exemple, ASER, Uwezo, etc.)	CLA	Variable selon le pays, habituellement de 5/6 à 16/18 ans	Compétences de base en lecture Compétences de base en mathématiques	Variables
	Évaluation de l'apprentissage précoce de la lecture	EGRA	Niveaux 1-3	Compétences de base en lecture	Variables
	Évaluation de l'apprentissage précoce des mathématiques	EGMA	Niveaux 1-3	Compétences de base en mathématiques	Variables

* À l'étape des essais sur le terrain. Des comparaisons internationales pourraient être établies à l'avenir.

Bien que la lecture/écriture et le calcul soient les domaines d'apprentissage les plus fréquemment évalués à l'échelle mondiale dans le cadre des évaluations nationales et internationales, un système solide d'évaluation de l'apprentissage inclurait la plupart des savoir-faire et des compétences prescrits dans le programme national. Dans la plupart des pays, le nombre de domaines d'apprentissage visés est largement supérieur à deux³³. L'initiative pour la mesure de l'apprentissage à l'école primaire en Asie du Sud-est (SEA-PLM), le Programme international pour le suivi des acquis des élèves (PISA) et l'Étude internationale sur l'éducation à la citoyenneté de l'Association internationale pour l'évaluation du rendement scolaire (IEA), entre autres, englobent des domaines liés à la citoyenneté mondiale et à la résolution de problèmes en collaboration. Le Consortium d'Afrique australe pour le suivi de la qualité de l'éducation (SEACMEQ) s'intéresse à la sensibilisation au VIH/SIDA et aux compétences nécessaires à la vie courante.

Plus de 80 % des pays du monde entier ont procédé à un type quelconque d'évaluation à grande échelle des compétences en lecture et en mathématiques entre 2010 et 2015³⁴. Des 65 pays en développement partenaires du PME pendant cette période, 50 (77 %) ont participé à un type quelconque d'évaluation nationale, régionale ou internationale. La prise en compte dans ce calcul des évaluations EGRA/EGMA et des évaluations citoyennes porte le nombre de PDP participants à l'un ou l'autre de ces types d'évaluations à 55 (85 %)³⁵. Nous présentons à l'annexe B une liste des PDP et des évaluations auxquelles ils ont participé.

Les évaluations des aptitudes à la lecture et aux mathématiques dans les petites classes sont celles auxquelles les PDP participent en plus grands nombres (41 pays, ou 63 %), bien que les échantillons retenus ne soient pas toujours représentatifs au plan national. Cependant, ces évaluations peuvent servir de point de départ pour des études plus vastes. Beaucoup de PDP réalisent aussi des évaluations nationales (39 pays, ou 60 %) et des évaluations régionales (25 pays, ou 38 %). Onze PDP réalisent des évaluations citoyennes et des évaluations internationales (14 % pour chaque type d'évaluation). Neuf PDP ne participent à aucune des évaluations susmentionnées depuis 2010 : Comores, Djibouti, Dominique, Grenade, Mauritanie, République centrafricaine, Sainte-Lucie, Saint-Vincent-et-les Grenadines et Ouzbékistan³⁶. Ce sont pour la plupart des petits États insulaires en développement (PEID) ou des pays touchés par la fragilité ou les conflits qui n'ont peut-être pas les ressources et les infrastructures requises pour élaborer des évaluations de l'apprentissage ou y participer.

Données

Le site Web de l'ISU est la source principale des données portant sur le secteur de l'éducation. Il comprend les données disponibles concernant l'ensemble des indicateurs du quatrième ODD. Les évaluations énumérées dans la section précédente offrent aux pays participants l'occasion de publier leurs résultats par l'intermédiaire d'une plateforme centrale en ligne, parfois sous la forme de rapports nationaux individuels, et parfois sous la forme de synthèses englobant les résultats de plusieurs pays et de rapports interactifs. Certaines plateformes en ligne comprennent aussi des ensembles de données. Par exemple, l'USAID a lancé en 2018 le *Early Grade Reading Barometer* qui publie des rapports et des

³³ Care, Esther, Helyn Kim, Kate Anderson et Emily Gustafsson-Wright. « *Skills for a Changing World: National Perspectives and the Global Movement* ». Washington DC: Brookings Institution, 2017.

³⁴ Institut de statistique de l'UNESCO, 2018.

³⁵ Les études EGRA et EGMA servent souvent à l'évaluation des programmes et ne portent donc pas toujours sur des échantillons représentatifs au plan national.

³⁶ L'Ouzbékistan compte participer à l'initiative PISA en 2021.

données sur les pays qui ont utilisé diverses versions des outils EGRA³⁷. Ce site Web affiche les données de diverses façons, y compris sous la forme d'une vue d'ensemble du niveau d'apprentissage de la lecture par les élèves d'un pays donné, du pourcentage des élèves qui éprouvent des difficultés en lecture, des résultats des élèves aux sous-tâches d'EGRA, et de comparaisons des résultats EGRA d'une région à l'autre. Le réseau PAL « L'action citoyenne pour l'apprentissage » propose des ensembles de données téléchargeables issues de plusieurs des pays participant aux évaluations citoyennes.

La base de données du catalogue des mesures de l'apprentissage de l'ISU, créée avec l'aide financière du PME, constitue une autre source de métadonnées utiles sur les évaluations de l'apprentissage. Elle fournit des informations sur divers aspects des évaluations de l'apprentissage représentatives au plan national menées dans les pays du monde entier. Le catalogue inclut les résultats d'évaluations réalisées en classe et dans les foyers. Il couvre les évaluations réalisées au niveau de la petite enfance et des cycles primaire et secondaire, et les groupes d'âge correspondant à ces niveaux (pour les évaluations réalisées auprès des ménages). L'Indice de la capacité à évaluer l'apprentissage (LACI) présente ces résultats sous la forme d'une carte mondiale. En utilisant cet outil et plusieurs autres sources de données, il est possible de déterminer les types d'évaluations auxquels les PDP participent (voir annexe B).

Réseaux et partage des connaissances

Divers forums permettent à ceux qui travaillent aux évaluations à grande échelle de communiquer entre eux et de partager leurs connaissances. Chacune des évaluations à grande échelle énumérées au tableau 2 comporte un type quelconque de création de réseaux prévoyant d'ordinaire des réunions en présentiel normales et des formes virtuelles de collaboration. L'IEA a intégré la plupart des données et documents provenant de l'ensemble des évaluations internationales dans une plateforme en ligne : le portail ILSA (*International Large-Scale Assessment*).

L'Alliance mondiale pour le suivi de l'apprentissage (GAML) est une plateforme institutionnelle mise en place par l'ISU pour harmoniser les normes et coordonner les efforts déployés mondialement pour mesurer l'apprentissage. Cette alliance rassemble des experts du monde entier afin d'élaborer des méthodes fondées sur des indicateurs d'apprentissage pour suivre les progrès vers la réalisation du quatrième ODD et pour établir des normes de bonnes pratiques d'évaluation de l'apprentissage. Bien que les produits de l'Alliance soient des biens mondiaux, ce réseau est aussi le seul à réunir la presque totalité des parties prenantes à l'évaluation de l'apprentissage dans le monde, et les activités de réseautage organisées par ses membres peuvent aussi être considérées comme un bien public. Le Portail mondial de l'IPE est une plateforme mondiale qui fournit une banque du savoir, diffuse des blogs et organise des webinaires sur la planification de l'apprentissage et de l'éducation, y compris plusieurs documents sur le suivi de l'apprentissage.

Des organisations professionnelles comme l'Association internationale pour l'évaluation éducative (AIEE) et le *National Council on Measurement in Education* (NCME) permettent aux chercheurs et aux fonctionnaires de partager leurs connaissances et de se pencher sur un large éventail d'enjeux liés à l'évaluation. Au niveau régional, le travail du NEQMAP et de TALENT (dans la région Asie-Pacifique et en Afrique subsaharienne respectivement) facilite le renforcement des capacités et le partage du savoir et produit des études internationales. Le Réseau pour l'évaluation de l'apprentissage en Afrique (NALA), mis en place par l'Association pour le développement de l'éducation en Afrique (ADEA) et les

³⁷ USAID. « *Early Grade Reading Barometer* ». Visité le 8 mai 2018. <http://www.earlygradereadingbarometer.org/>.

gouvernements du Kenya, du Rwanda, du Sénégal et de la Zambie pour faciliter l'apprentissage entre pairs sur l'évaluation dans la région, constitue un autre exemple pertinent³⁸. Des organisations professionnelles régionales comme l'Association pour l'évaluation éducative en Afrique (AEEA) offrent aussi des plateformes de partage régional des connaissances, spécifiquement entre les organismes examinateurs.

Initiatives de renforcement des capacités

Une deuxième catégorie de biens mondiaux couvre les efforts déployés pour renforcer les capacités en matière d'évaluation de l'apprentissage. Elle inclut les outils et les publications qui aident les pays à évaluer l'apprentissage, les informations sur les capacités du système, et les efforts en matière de partage des connaissances.

Outils et publications

Les cinq volumes de la série de rapports sur les évaluations nationales du rendement scolaire publiés par la Banque mondiale de 2008 à 2015 fournissent des directives générales sur la mise en place d'un système national d'évaluation, l'élaboration et l'utilisation de tests et de questionnaires, la mise en œuvre des systèmes d'évaluation, l'utilisation de l'information et l'analyse des données. Ces documents constituent la ressource la plus complète d'informations sur l'évaluation nationale des résultats éducatifs. Ils sont disponibles en ligne, et la Banque mondiale en fournit sur demande sans frais des versions imprimées, dont certains s'accompagnent de CD-ROM contenant des ressources supplémentaires³⁹. Ces documents produits en anglais ont été traduits en plusieurs langues (arabe, français, portugais, russe et espagnol). La Banque mondiale offre aussi une formation en ligne sur l'évaluation des acquis scolaires. Ce cours d'introduction succinct sur l'évaluation de l'apprentissage a été élaboré spécialement pour les décideurs et les praticiens des pays en développement⁴⁰. L'IIPE de l'UNESCO propose par ailleurs une série de publications en dix volumes portant sur les méthodes de recherche quantitative dans le domaine de la planification de l'éducation. Ces documents couvrent entre autres la recherche et les plans de sondage, la rédaction et l'analyse des questions et la mise à l'essai des tests. Cette série comprend par ailleurs des cours en ligne ouverts à tous (MOOC) sur l'évaluation de l'apprentissage.

Pour améliorer la pertinence et la cohésion du système d'évaluation, l'Alliance mondiale pour le suivi de l'apprentissage (GAML) a élaboré un ensemble de « principes » sur les bonnes pratiques en matière d'évaluation de l'apprentissage (GP-LA) qui expliquent comment élaborer et mettre en œuvre de solides programmes d'évaluation à grande échelle et faire en sorte que les données puissent servir au suivi des systèmes éducatifs et à l'élaboration de politiques de l'éducation fondées sur des données probantes. Les GP-LA établissent des principes pertinents aux diverses activités d'évaluation à grande échelle de l'apprentissage menées à travers le monde, y compris les évaluations internationales et nationales. L'Alliance publie également à l'intention des pays une série de petits guides contenant des instructions détaillées sur la préparation de rapports sur les indicateurs du quatrième ODD⁴¹.

³⁸ Mugo, John. « *Welcoming Two New Initiatives to Accelerate Quality Teaching and Learning in Africa* ». Washington DC: Brookings Institution, 2017

³⁹ Greaney, Vincent et Thomas Kellaghan. 2008. « *Assessing National Achievement Levels in Education* ». Washington, DC: Banque mondiale, 2008.

⁴⁰ Ces documents sont disponibles (en anglais) à l'adresse suivante : <https://openknowledge.worldbank.org/handle/10986/2143>.

⁴¹ Ramirez, Maria Jose. « *Quick Guide No. 3 Implementing a National Learning Assessment* ». Montréal: UIS, 2017.

Le Fonds fiduciaire de l'assistance russe à l'éducation pour le développement (READ) de la Banque mondiale est un autre exemple d'outil de renforcement des capacités nationales qui a aussi produit des biens publics mondiaux. Depuis 2008, ce fonds a fourni environ 38 millions de dollars à l'appui des systèmes nationaux d'évaluation. Les leçons tirées de ces expériences ont servi à élaborer une série de supports mondiaux de connaissances sur les évaluations.

Informations sur les capacités des systèmes

Le programme SABER d'évaluation des élèves de la Banque mondiale fournit des ressources et des données sur l'évaluation de la qualité globale des systèmes d'évaluation de l'éducation. Il propose une trousse d'instruments de diagnostic constituée de questionnaires normalisés et de rubriques qui permettent aux utilisateurs de mesurer les caractéristiques des outils d'évaluation de l'apprentissage à l'aune de normes de qualité applicables à tous les types d'activités d'évaluation entreprises par les administrations publiques, y compris les examens nationaux, les évaluations nationales basées sur des échantillons, les évaluations internationales et régionales et les évaluations en classe. L'Indice de la capacité à évaluer l'apprentissage (LACI) de l'ISU propose une carte des résultats obtenus par les pays à divers niveaux de capacité d'évaluation à grande échelle de l'apprentissage fondée sur la participation de ces pays aux évaluations nationales, régionales et internationales⁴².

Le PME s'emploie également à élaborer un outil complémentaire d'aide au diagnostic des systèmes d'évaluation de l'apprentissage connu sous le nom d'Analyse des systèmes nationaux d'évaluation des acquis (ANLAS). Cet outil vise en particulier à étayer les stratégies de renforcement de ces systèmes ; il doit être mis en œuvre dans le cadre des plans sectoriels de l'éducation nationaux. L'ANLAS est le premier outil de diagnostic intégrant une analyse de la mesure dans laquelle les systèmes d'évaluation de l'apprentissage prennent en compte les compétences du XXI^e siècle. Il fera l'objet d'essais pilotes dans trois PDP en 2019, et sera ensuite finalisé et diffusé.

Réseaux et partage des connaissances

Les évaluations internationales et régionales offrent aux pays participants un volet de renforcement des capacités⁴³. L'évaluation des aptitudes à la lecture, à l'écriture et au calcul élaborée par l'IEA aide les pays en développement à renforcer leurs capacités en matière d'évaluation de l'apprentissage à l'échelle locale et appuie les efforts déployés par la communauté éducative internationale dans ce domaine⁴⁴. L'Organisation de coopération et de développement économiques (OCDE) a aussi élaboré des ressources pour le renforcement des capacités des pays en développement. Les pays qui participent au programme PISA-D et certains de ceux qui participent au PISA ont réalisé une analyse des besoins en matière de capacités fondée sur les rubriques du programme SABER d'évaluation des élèves et sur les normes de qualité de l'OCDE. Ces pays ont aussi collaboré avec l'OCDE à l'élaboration d'un plan de renforcement des capacités visant à appuyer leur participation fructueuse à l'évaluation PISA-D⁴⁵. À ce jour, de telles analyses ont été réalisées ou de tels plans ont été élaborés pour le Cambodge, l'Équateur, le Guatemala, le Honduras, le Panama, le Paraguay, le Sénégal, la Tanzanie, l'Ukraine et la Zambie. Malgré leurs promesses, ces efforts sont liés à des études particulières. Il conviendra à l'avenir

⁴² Institut de statistique de l'UNESCO, 2018.

⁴³ Lockheed, Marlaïne E. « *The Craft of Education Assessment: Does Participating in International and Regional Assessments Build Assessment Capacity in Developing Countries? An Independent Evaluation of IEA's Program on the Assessment of Student Achievement (PASA)* », 2010.

⁴⁴ IEA. « IEA Studies ». Consulté le 8 mai 2018.

⁴⁵ OCDE. « *Building Capacity for Assessment in PISA for Development Countries* ». Paris: OECD, 2017.

d'intensifier le réseautage et le partage des connaissances dans le domaine du renforcement des capacités sans égard au choix des pays de participer ou non aux évaluations internationales.

Innovations relatives à l'évaluation de l'apprentissage

La troisième catégorie de biens mondiaux est celle des initiatives qui visent à innover dans le domaine de l'évaluation de l'apprentissage.

Outils

L'ISU examine actuellement des moyens de comparer les résultats d'apprentissage entre des pays utilisant différentes évaluations de l'apprentissage aux fins de l'établissement des rapports sur le quatrième ODD. La première étape de ce processus consistait à élaborer des échelles de rapport communes en collaboration avec l'*Australian Council for Educational Research* (ACER). Ces échelles numériques s'accompagnent de descriptions explicatives détaillées de l'évolution des niveaux de maîtrise des matières dans les domaines d'apprentissage définis dans le cadre des indicateurs de résultats de l'ODD. Par exemple, pour l'indicateur 4.1.1 sur les niveaux minimaux pour la maîtrise de la lecture et des mathématiques, les descriptions détaillées proposent un ensemble à complexité croissante d'aptitudes en lecture et en mathématiques. Certains points particuliers sur les échelles de rapport de l'ISU deviendraient des jalons permettant de formuler des définitions générales communes de notions telles que « compétences minimales » et « en bonne voie de développement ». L'explorateur de la progression de l'apprentissage est un outil disponible en ligne qui permettra aux utilisateurs d'étudier ces échelles de rapport. Cette innovation constitue une nouvelle façon de mettre en équation les résultats d'apprentissage entre divers pays et les parcours éducatifs des apprenants, de la petite enfance jusqu'au niveau secondaire.

Le processus d'élaboration des échelles de rapport a suscité des débats sur la question de savoir s'il était possible d'élaborer une échelle de comparaison véritablement universelle entre des pays, des langues et des contextes multiples. C'est la raison pour laquelle l'Alliance mondiale pour le suivi de l'apprentissage se penche également sur une approche de « soutien social » élaborée par *Management Systems International* (MSI). En vertu de cette approche, des groupes d'experts nationaux, y compris des enseignants, examinent les résultats des évaluations nationales et internationales pour déterminer dans quelle mesure ils correspondent aux descriptions des niveaux de compétence.

Recherche

Les biens mondiaux applicables aux évaluations à l'échelle de la classe sont plus rares puisque la pratique pédagogique et l'évaluation en classe sont souvent considérées comme l'affaire des enseignants et des dirigeants d'écoles, et qu'elles dépendent étroitement du programme scolaire et des besoins du système national d'éducation en matière d'instruction. Il existe toutefois des exemples d'évaluations simples utilisées pour recueillir des informations sur les apprenants et pour planifier des interventions qui ont entraîné une amélioration des acquis scolaires. Par exemple, le Read India program mis en œuvre par l'ONG Pratham utilise l'évaluation des aptitudes à la lecture, à l'écriture et au calcul du rapport annuel sur la situation de l'éducation (ASER) pour classer les enfants par niveaux d'aptitude. Un enseignement adapté au niveau individuel d'apprentissage a donné chez certains enfants des résultats meilleurs et durables⁴⁶. Ce programme a été étendu au cours des récentes années en

⁴⁶ Banerji et Walton, 2011.

fournissant aux élèves des tablettes proposant des jeux de lecture et de mathématiques et favorisant la participation des parents⁴⁷.

Il existe de nombreux exemples d'utilisation par les pays des données des évaluations à grande échelle pour étayer les politiques. Best et ses collaborateurs (2013) ont constaté que les résultats des évaluations à grande échelle servent le plus souvent à la formulation de politiques sur l'allocation des ressources, les normes et les réformes académiques, l'évaluation des performances, la normalisation et l'évaluation. Les pratiques d'enseignement et d'apprentissage, les politiques favorables à l'engagement de la collectivité et des parents et la responsabilisation sont les enjeux stratégiques au sujet desquels les données sont le moins souvent utilisées. La qualité insuffisante des programmes d'évaluation et l'absence d'analyses approfondies des données comptent au nombre des obstacles qui entravent le plus souvent l'utilisation des données d'évaluation. Dans la région Asie-Pacifique, les évaluations sont d'abord utilisées par les décideurs pour le suivi et l'évaluation des politiques d'éducation, mais aussi pour éclairer la mise en œuvre des politiques, la définition des programmes et la formulation de nouvelles politiques⁴⁸.

Le projet *Optimizing Assessment for All* (OAA) lancé en 2017 par la Brookings Institution en collaboration avec le PME est un exemple d'initiative conçue pour appliquer les bonnes pratiques mondiales au renforcement des capacités en matière d'évaluations en classe. Le projet met l'accent sur les régions Afrique et Asie et collabore avec les réseaux régionaux et les pouvoirs publics à l'élaboration d'évaluations adaptées à la promotion des compétences du XXI^e siècle⁴⁹. L'éventail des initiatives offertes par l'Internationale de l'Éducation et la Brookings Institution procure des outils utiles pour mesurer l'alignement des opportunités d'apprentissage vues sous l'angle des enseignants et des administrateurs scolaires, ainsi qu'au niveau des politiques⁵⁰. Ces outils à code source libre comprennent des questionnaires sur l'évaluation en classe, des examens nationaux et des évaluations à grande échelle auxquelles les pays peuvent participer.

6. Carences des biens mondiaux disponibles

La plupart des biens mondiaux existants sont axés sur les points suivants : 1) appui aux groupes de pays pour la conception et l'administration d'évaluations à grande échelle de l'apprentissage à caractère particulier ; 2) regroupement des réseaux de professionnels intéressés aux enjeux de l'évaluation ; 3) fourniture de ressources en ligne et de directives générales sur la façon d'élaborer des évaluations de l'apprentissage de qualité. Les parties prenantes qui ont participé aux consultations menées dans le cadre de la présente étude ont généralement estimé que la section 5 proposait un résumé détaillé des biens mondiaux disponibles pour le renforcement des systèmes d'évaluation de l'apprentissage.

La perception des lacunes des biens mondiaux varient selon le groupe concerné. Les PDP ont jugé, à une écrasante majorité, qu'il existait des lacunes dans la formation des enseignants à l'utilisation des

⁴⁷ Winthrop, Rebecca, Eileen McGivney et Adam Barton. « *Can We Leapfrog? The Potential of Education Innovations to Rapidly Accelerate Progress* ». Washington DC: Brookings Institution, 2017.

⁴⁸ Tobin, Mollie, Petra Lietz, Dita Nugroho, Ramya Vivekanandan et Tserennadmid Nyamkhuu. « *Using Large-scale Assessments of Students Learning to Inform Education Policy: Insights from the Asia-Pacific Region* ». Melbourne et Bangkok : ACER et UNESCO Bangkok, 2015.

⁴⁹ Care, Vista et Kim, 2018.

⁵⁰ Anderson *et al.*, 2018.

évaluations en classe, ainsi que dans les bonnes pratiques en matière de réforme des examens. Ils ont aussi déploré un manque de directives sur la façon d'organiser les données issues des évaluations de l'apprentissage de manière à ce qu'elles soient utiles aux décideurs, aux enseignants et aux autres parties prenantes. Bien que la participation aux évaluations à grande échelle assure dans une certaine mesure le renforcement des capacités de certains agents ministériels, les PDP ont estimé qu'il convenait d'étendre cette participation pour établir une culture globale de l'évaluation. Les PDP ont aussi relevé des incohérences entre les divers systèmes d'évaluation de l'apprentissage, et demandé des conseils sur la façon d'établir de solides institutions pour le suivi de l'ensemble du portefeuille national d'évaluation.

Les experts internationaux s'inquiètent davantage de l'absence de données comparables sur les acquis scolaires au niveau international, et ont dit souhaiter que plus de pays participent aux évaluations internationales. Plusieurs ont noté la nécessité d'accorder à certaines initiatives régionales d'évaluation un appui budgétaire supplémentaire.

7. Domaines d'investissement envisageables

Le processus de recherche et de consultations a débouché sur un certain nombre de possibilités au sujet desquelles les biens mondiaux financés par le mécanisme KIX pourraient jouer un rôle crucial. Ces possibilités s'harmonisent également avec le Cadre de résultats et le Modèle de financement du PME. Elles sont examinées ci-dessous.

Première possibilité : Biens mondiaux appuyant les institutions nationales d'évaluation de l'apprentissage.

Ces biens mondiaux incluent les ministères de l'éducation et les agences autonomes et semi-autonomes d'évaluation. Nous avons par exemple besoin d'activités de renforcement des capacités pour aider les pays à établir des cadres nationaux d'évaluation. Les participants à la consultation ont confirmé qu'un cadre d'évaluation solide 1) permet d'améliorer l'apprentissage et les acquis scolaires grâce à un recours efficace aux évaluations en classe qui permettent d'améliorer la pédagogie ; 2) fait en sorte que les examens, lorsqu'ils sont utilisés, certifient l'acquisition des connaissances ; 3) favorise l'utilisation des évaluations à grande échelle pour surveiller et appuyer les systèmes et les processus visant à améliorer les conditions d'enseignement et d'apprentissage⁵¹.

L'échange régional d'informations sur les moyens d'établir de solides institutions d'évaluation compte parmi les moyens de renforcer les capacités dans ce domaine. De nombreuses initiatives régionales de partage des connaissances propices à un tel échange sont déjà en cours. Elles pourraient aussi fournir l'occasion d'étendre la portée des outils de diagnostic des systèmes d'évaluation de l'apprentissage et servir à recommander des améliorations — par exemple, le cadre de suivi des acquis des élèves SABER et l'outil d'analyse des systèmes nationaux d'évaluation des acquis (ANLAS) du PME.

L'orientation des pratiques concernant la réforme des systèmes d'examen nationaux constitue une possibilité majeure pour le mécanisme KIX. Dans beaucoup de pays, les examens sont de piètre qualité et font obstacle aux efforts déployés dans d'autres domaines d'évaluation⁵². Dans certains systèmes éducatifs, le nombre de places limité en éducation secondaire et tertiaire pousse les autorités à utiliser

⁵¹ Kanjee, Anil. Extrait d'un rapport de consultation. 10 août 2018.

⁵² Comité de réflexion sur la métrique des apprentissages, 2014 ; Burdett, Newman. « *Review of High Stakes Examination Instruments in Primary and Secondary School in Developing Countries* ». RISE Working Paper 17/018, 2017.

les examens publics pour exclure une partie des élèves des niveaux supérieurs d'éducation. En revanche, les examens publics placent la barre trop bas pour le niveau d'instruction que les apprenants sont censés atteindre à la fin de leurs études. Par exemple, une étude réalisée en 2017 sur les examens administrés en Ouganda, au Nigéria, au Pakistan et en Inde a observé que la plupart des questions posées portaient sur des compétences d'ordre inférieur comme la mémorisation de faits, et que quelques-unes seulement testaient des compétences d'ordre supérieur comme le raisonnement critique⁵³. Bien que des preuves corrélationnelles donnent à penser que les élèves de pays administrant des examens au deuxième cycle du secondaire réussissent mieux aux évaluations internationales comme TIMSS et PISA, il convient d'examiner plus avant les liens qui peuvent exister entre les examens et les niveaux d'apprentissage, en particulier dans les pays en développement⁵⁴.

Les examens nationaux ont tendance à accaparer la majorité des ressources internes consacrées à l'évaluation de l'éducation puisqu'ils sont administrés chaque année, visent la population entière des élèves, et sont souvent élaborés, administrés et certifiés par des organisations publiques ou quasi-publiques mises en place à cette fin. On reproche le plus souvent aux systèmes d'examen nationaux de nombreux pays en développement d'influer sur les pratiques des enseignants et de les inciter à ne mettre l'accent que sur les matières du programme scolaire qui feront l'objet des examens.

Les examens ont pour but de certifier la conclusion satisfaisante de la scolarité et fournissent une mesure de l'engagement des autorités publiques à fournir une éducation de qualité à l'ensemble des citoyens. Ils permettent aussi d'accéder aux études supérieures et à l'emploi, et d'échapper à la pauvreté. En raison de leur importance, la tricherie, la fraude et les malversations sont monnaie courante dans les systèmes d'examen de certains PDP⁵⁵. La falsification des certificats est fréquente, et les systèmes d'examen sont minés par la corruption⁵⁶. Les élèves ont souvent le sentiment que leurs résultats d'examens ne reflètent pas fidèlement leurs connaissances et leurs compétences. En outre, les employeurs ont de plus en plus tendance à court-circuiter le système d'examen dans certains secteurs comme celui des technologies de l'information, lorsqu'ils peuvent compter sur des plateformes en ligne de reconnaissance des compétences proposées par des tiers⁵⁷. Cela dit, l'idée de supprimer ou même de réformer les systèmes d'examen nationaux est politiquement irréalisable dans la plupart des pays puisque plusieurs en tirent parti. Dans certains pays très corrompus, les systèmes d'examen, bien que corrompus, le sont malgré tout moins que les autres institutions et les pouvoirs publics répugnent à se départir de ces « îlots » d'intégrité relative. La collecte de données probantes sur les solutions de rechange aux systèmes d'examen périmés est un domaine où des biens mondiaux sont requis.

D'autres domaines, également jugés importants sur le plan de la vie professionnelle et personnelle, sont pourtant sous-évalués dans les systèmes d'éducation. Bien que les évaluations de l'apprentissage aient tendance à mettre l'accent sur la lecture et le calcul, le programme scolaire de la plupart des pays est beaucoup plus vaste⁵⁸. Parmi les domaines importants, mais pourtant sous-évalués, on peut mentionner

⁵³ Burdett, 2017.

⁵⁴ Clarke, 2012.

⁵⁵ Petters, Janet Sunday et Maureen O. Okon. « Students' Perception of Causes and Effects of Examination Malpractice in the Nigerian Educational System: The Way Forward for Quality Education ». *Procedia - Social and Behavioral Sciences* 114: 125-29, 2014.

⁵⁶ Transparency International. « *Global Corruption Report: Education* ». London: Earthscan, Publishing for a Sustainable Future, Routledge, Taylor & Francis Group, 2013 ; Mutambo, Aggrey. « UK Company Fined Sh330 Million for Bribing Kenya Officials ». *Daily Nation*. 11 janvier 2016.

⁵⁷ Winthrop, McGivney et Barton, 2017.

⁵⁸ Care *et al.*, 2017.

les compétences du XXI^e siècle, le développement social et affectif, et la citoyenneté mondiale. On s'emploie à élaborer des outils pour l'évaluation de l'apprentissage dans ces domaines ; le travail est plus avancé dans certains d'entre eux, mais d'autres posent des difficultés inhérentes et d'autres encore font l'objet d'un débat sur la valeur de précision et l'utilité des évaluations à grande échelle. Il s'agit entre autres des compétences du XXI^e siècle comme l'esprit critique, la collaboration, la résolution de problèmes et la créativité.

À ce propos, des changements stratégiques, en cours dans beaucoup de pays, mettent l'accent sur les programmes axés sur les compétences et sur l'évaluation de l'apprentissage. L'expression « évaluations axées sur les compétences » est largement utilisée, mais elle fait l'objet de diverses interprétations. Un examen des bonnes pratiques d'évaluation axée sur les compétences pourrait constituer un bien mondial utile pour beaucoup de PDP engagés dans la mise en place d'un programme d'enseignement axé sur les compétences.

Deuxième possibilité : Bien mondiaux qui renforcent les compétences en matière d'évaluation de l'apprentissage à tous les niveaux du système éducatif, des enseignants aux hauts responsables gouvernementaux.

La difficulté réside non pas dans le manque de savoir-faire, mais plutôt dans le manque de volonté politique et de financements pour mettre à l'échelle les solutions gagnantes. Des directives adaptées de renforcement des capacités techniques des responsables ministériels incluant une assistance technique en temps réel lorsqu'elle est nécessaire constitueraient un utile bien mondial.

Les innovations peu coûteuses dans le domaine du renforcement des compétences des responsables ministériels et des enseignants pourraient s'avérer utiles. Elles pourraient inclure la création de biens publics mondiaux en vertu desquels l'apprentissage dans un pays donné est conçu dès le départ pour bénéficier à l'apprentissage dans les autres pays.

Nous avons besoin de plus de données pour renforcer les capacités des enseignants en matière d'évaluations en classe, y compris celles réalisées par les enseignants, et celles réalisées par les apprenants (autoévaluations et portefeuilles). Lorsqu'elles sont réalisées fréquemment (au moins chaque semaine), les évaluations en classe peuvent améliorer l'apprentissage⁵⁹. Les PDP semblent intéressés à renforcer les capacités des chefs d'établissement et des enseignants à faire usage de l'évaluation continue pour améliorer l'apprentissage. Par exemple, l'initiative *Learning Champions* de la Brookings Institution a laissé constater un intérêt considérable de la part des autorités publiques et de la société civile pour l'évaluation continue en Éthiopie, au Pakistan, en Palestine, au Sénégal, en Tunisie et en Zambie, en même temps qu'elle constatait un manque de ressources chez les fonctionnaires cherchant à appuyer la réalisation de ces évaluations⁶⁰. Il convient d'accroître la masse des connaissances disponibles sur la façon dont les enseignants peuvent être le mieux préparés pour utiliser l'évaluation en classe afin d'accroître l'apprentissage, en particulier dans le contexte des PDP.

Donner aux systèmes d'éducation les moyens d'élaborer de nouvelles méthodes d'évaluation pour l'ensemble du programme scolaire est un autre domaine dans lequel des avancées seraient propices. Les innovations à ce niveau porteraient à la fois sur les aptitudes académiques traditionnelles et sur les

⁵⁹ Black, Paul et Dylan Wiliam. « Developing a Theory of Formative Assessment ». In J. Gardner, John, éd. *Assessment and Learning*. London: SAGE, 2006.

⁶⁰ Anderson, Kate et Joshua Muskin. « *Learning Champions: How 15 Countries, Cities, and Provinces Came Together to Rethink Learning Assessment* ». Washington DC: Brookings Institution, 2018.

domaines sous-évalués — par exemple, mais sans s'y limiter, la réflexion critique, la résolution de problèmes et la créativité.

Troisième possibilité : Bien mondiaux appuyant les systèmes d'évaluation de l'apprentissage des élèves les plus marginalisés.

Les systèmes d'évaluation de l'apprentissage sont lacunaires dans la plupart des pays les plus marginalisés, y compris les PSC, et les apprenants les plus marginalisés des différents pays. Dans la plupart des pays vulnérables, les données représentatives au niveau national sont insuffisantes pour appuyer les interventions qui améliorent l'apprentissage⁶¹. Des biens mondiaux sont requis pour aider les PSC qui n'effectuent pas d'évaluations à participer aux évaluations internationales répondant aux normes de qualité du PME. Ces biens mondiaux pourraient aussi inclure des examens de la rentabilité des évaluations et des solutions qui permettent aux PSC de participer aux évaluations nationales et internationales de l'apprentissage. Plusieurs choix s'offrent aux pays lorsqu'il s'agit d'évaluer les acquis de l'apprentissage, mais on connaît mal la rentabilité relative des diverses méthodes d'évaluation. Pour les pays qui ne disposent pas d'un système d'évaluation de l'apprentissage ou pour ceux où l'apprentissage et son évaluation ont été interrompus par un conflit ou par la fragilité, il importe de savoir quel type d'évaluation fournira les informations les plus pertinentes au coût le plus bas. De plus amples études sont requises sur les meilleures façons de choisir les options d'évaluation dans les situations caractérisées par un manque de financements, de ressources humaines et de capacités.

Les biens mondiaux actuels pour l'évaluation de l'apprentissage excluent généralement certains groupes, comme les enfants vivant avec un handicap et les populations non scolarisées ou déplacées. Il est nécessaire d'innover pour mesurer les acquis scolaires de tous les enfants. Nombre des parties prenantes consultées dans le cadre de la présente étude ont fait état des difficultés rencontrées par les PSC qui cherchent à rationaliser les priorités des donateurs pour établir un système d'évaluation répondant à leurs besoins. Des directives sur l'harmonisation des priorités des donateurs pourraient constituer un bien mondial utile.

L'ISU et l'ACER font des progrès notables en matière d'harmonisation des évaluations internationales et régionales. Cependant, il n'existe toujours pas de base de référence comparative mondiale pour les résultats d'apprentissage qui permettrait de suivre les progrès vers la réalisation du quatrième ODD. Cela limite les possibilités de comparaisons transnationales et d'apprentissage par les pairs, en particulier dans les pays en développement. L'établissement d'une base de référence mondiale nécessitera un effort collectif de la part des pouvoirs publics nationaux, des initiatives régionales d'évaluation, et des organisations internationales. D'autres innovations sont requises dans ce domaine, notamment en ce qui a trait à la façon d'obtenir des données auprès du plus grand nombre possible de pays sans surcharger les systèmes nationaux et détourner leurs maigres ressources pour réaliser des évaluations dont ils n'auront pas l'occasion d'utiliser les résultats. Les investissements requis à cette fin pourraient aussi financer des mécanismes d'aide à l'utilisation des données existantes aux fins de l'établissement de rapports au niveau mondial, et contribuer à l'élaboration de solutions technologiques peu coûteuses pour la collecte de données dans les PSC et auprès des élèves handicapés ou déplacés.

⁶¹ ISU, 2017a.

Quatrième possibilité : Biens mondiaux à l'appui de l'utilisation systémique des données d'évaluation de l'apprentissage.

Cet enjeu recoupe en partie celui des travaux thématiques sur les données du KIX, mais les PDP et les experts consultés se sont dits convaincus qu'il devrait être pris en compte. Les échanges régionaux et internationaux sur les bonnes pratiques de présentation et d'utilisation des données d'évaluation des acquis de l'apprentissage pourraient être utiles aux PDP. Un examen de la manière dont les données d'évaluation de l'apprentissage (évaluations en classe, examen et évaluations à grande échelle) influent sur les politiques et les pratiques d'enseignement et d'apprentissage dans les pays à revenu faible ou moyen pourrait aussi faire œuvre utile.

Certains pays ont élaboré des moyens de dériver l'information des examens et de l'utiliser pour améliorer l'enseignement en classe. Le Kenya fait partie du groupe⁶². Comme les examens accaparent une part considérable des ressources consacrées à l'évaluation, il conviendra de consacrer beaucoup plus d'efforts au renforcement de ces systèmes et d'autres types d'évaluations axées sur la certification, la promotion et les transferts, en insistant sur l'aide qui permettra aux pays de faire un meilleur usage de ces mécanismes.

Les évaluations à grande échelle peuvent servir à améliorer l'apprentissage en classe, mais elles sont rarement utilisées à cette fin. Une étude réalisée dans 73 pays en développement sur l'utilisation des résultats des évaluations nationales, régionales et internationales à grande échelle a laissé constater que seuls quelques-uns de ces pays utilisaient ces résultats pour améliorer les pratiques d'enseignement et d'apprentissage comme les stratégies d'apprentissage et la pédagogie axée sur l'élève⁶³. Ce domaine en est un autre dans lequel les biens mondiaux pourraient conjuguer les recherches existantes pour constituer une base de données factuelles utiles pour les décideurs et les praticiens des PDP.

Les évaluations à grande échelle sont souvent mal utilisées puisque les évaluations nationales, en particulier, peuvent être de piètre qualité et ne se prêtent pas toujours très bien aux comparaisons sur plusieurs années. Les administrations publiques et les bailleurs cherchent parfois à élaborer et à mettre en œuvre de nouveaux outils d'évaluation sans examiner les moyens de « réutiliser » les résultats des évaluations existantes. Il convient de recueillir de plus amples informations sur la façon d'effectuer des analyses différentes et de créer des méthodes novatrices d'utilisation des données d'évaluation de l'apprentissage à de multiples fins.

Les évaluations à grande échelle peuvent promouvoir l'apprentissage en déterminant les domaines où les élèves ont du mal à suivre le rythme, en attirant l'attention sur ces domaines et en y consacrant plus de ressources⁶⁴. Elles peuvent aussi servir à évaluer et à modifier les politiques⁶⁵. Par exemple, le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) a permis d'attirer l'attention sur les problèmes de persévérance scolaire au Sénégal et de démontrer que le redoublement ne permettait pas d'améliorer les acquis scolaires l'année suivante. À Madagascar, les résultats du PASEC ont laissé

⁶² Somerset, H.C.A. « *Examination Reform in Kenya* ». Washington DC: Banque mondiale, 1987.

⁶³ Best, Maura, Pat Knight, Petra Leitz, Craig Lockwood, Dita Nugroho et Mollie Tobin. « *The Impact of National and International Assessment Programmes on Education Policy, Particularly Policies regarding Resource Allocation and Teaching and Learning Practices in Developing Countries* ». Melbourne: ACER, 2013.

⁶⁴ Kelleghan, Thomas, Vincent Greaney et T. Scott Murray. « *Using the Results of a National Assessment of Educational Achievement* ». Vol. 5. Washington, DC: Banque mondiale, 2009.

⁶⁵ Institut de statistique de l'UNESCO. « *Indice de la capacité à évaluer l'apprentissage* ». Montréal: UNESCO, ISU. Consulté le 8 mai 2018.

conclure à l'existence d'un lien entre la formation continue des enseignants et une détérioration des acquis scolaires. Une étude plus approfondie a révélé que ce problème était dû au fait que la formation était offerte pendant les heures de classe, réduisant ainsi le temps d'enseignement. Cette politique a donc été modifiée⁶⁶.

Bien qu'il apparaisse que les autorités publiques s'appuient sur les résultats des évaluations à grande échelle pour modifier leurs politiques, les études conduisent à douter de l'incidence de ces évaluations. Par exemple, les résultats d'une enquête menée en 2017 donnent à penser que les résultats des évaluations internationales et régionales servent plus souvent à légitimer les réformes déjà engagées⁶⁷. Il convient d'examiner plus en détail les liens entre la participation aux évaluations à grande échelle et les répercussions de cette participation sur les politiques et les pratiques des PDP. Bien qu'il existe des anecdotes sur les façons d'utiliser les évaluations à grande échelle pour améliorer les acquis de l'apprentissage, les données, telles qu'elles se présentent actuellement, sont inutiles pour les décideurs. La publication de conseils et d'exemples sur la façon d'utiliser les divers types d'évaluations pour améliorer les acquis l'apprentissage en classe permettrait de combler une grave carence de biens mondiaux. Les parties prenantes consultées dans le cadre de la présente étude ont laissé entendre que l'apprentissage entre pairs sur les façons d'utiliser les résultats des évaluations de l'apprentissage pourrait être particulièrement utile pour les décideurs. Un groupe d'experts internationaux mettant ses membres à disposition pour fournir une assistance technique sur l'utilisation des données aux fins de la prise de décisions pourrait l'être tout autant.

Enfin, il est également nécessaire d'innover pour promouvoir une utilisation systématique des données d'évaluation de l'apprentissage afin d'évaluer les interventions et les réformes. À cette fin, le recours à une approche « tester-intervenir-tester à nouveau » systématique fondée sur les évaluations des acquis pourrait jeter un éclairage sur la façon dont les pays peuvent incorporer ces évaluations dans leurs programmes d'éducation, et définir de nouvelles méthodes de communication et de visualisation pour améliorer l'utilisation des données sur l'évaluation des acquis. Une telle approche pourrait être expérimentée et mise à l'essai dans un petit nombre de pays, et ses résultats diffusés ensuite en tant que bien mondial.

Possibilités du mécanisme KIX et modèle du PME

Les recommandations du présent document sont liées de deux façons au Plan stratégique et au Cadre de résultats du PME. D'abord, elles visent à répondre aux besoins les plus importants définis par les PDP et les membres du partenariat du PME pour l'atteinte de résultats d'apprentissage meilleurs et plus équitables (but stratégique 1). Le PME, en sa qualité de partenariat composé de multiples parties prenantes, est en effet particulièrement bien placé faire le point sur les besoins dans ce domaine thématique et créer l'espace requis pour leur résolution. On constate donc sans surprise que les thèmes qui se dégagent du processus de consultation recourent aussi plusieurs des indicateurs du Cadre de résultats.

⁶⁶ Bernard, Jean-Marc et Katharina Michaelowa. « How Can Countries Use Cross-national Research Results to Address 'the Big Policy Issues?' » (Études de cas en Afrique francophone). In *Cross-national Studies of the Quality of Education: Planning Their Design and Managing Their Impact*, 229-40. Paris: IIPÉ, UNESCO, 2006.

⁶⁷ Fischman, Gustavo E., Amelia Marcetti Topper, Iveta Silova, Janna Goebel et Jessica L. Holloway. « An Examination of the Influence of International Large Scale Assessments and Global Learning Metrics on National School Reform Policies ». *Journal of Education Policy*, 7 mai 2018.

Le manque de moyens à l'échelle nationale est un problème majeur pour de nombreux pays et un obstacle au suivi des résultats d'apprentissage, une condition essentielle pour assurer l'efficacité du modèle de financement du PME. Le PME exige des PDP qu'ils mettent en place un système ou mécanisme de suivi des acquis scolaires ou qu'ils se dotent d'un plan d'élaboration d'un tel mécanisme assorti d'échéances précises. Les pays doivent pour cela disposer d'un système d'évaluation de l'apprentissage jugé de bonne qualité, tel que défini par l'indicateur 15 du Cadre de résultats (proportion de PDP dont le système d'évaluation des acquis scolaires dans l'éducation de base répond à des normes de qualité). Tous les domaines d'intervention prometteurs du mécanisme KIX abordés dans la présente section peuvent contribuer à la réalisation de cet indicateur et veiller ainsi à ce que les PDP puissent améliorer les évaluations de l'apprentissage. L'indicateur 1 du Cadre de résultats concernant l'amélioration des acquis de l'apprentissage ne saurait être respecté sans l'intervention de systèmes capables de générer les données requises sur l'apprentissage. Autrement dit, l'apprentissage ne saurait être amélioré sans une mesure adéquate des progrès et des régressions. Les domaines d'intervention du mécanisme KIX visent à mettre en place ou à renforcer les systèmes qui permettront d'évaluer cet apprentissage.

Pour veiller à ce que les investissements du mécanisme KIX destinés à ces domaines d'intervention répondent aux besoins et au contexte actuels, trois domaines d'investissement doivent être définis :

- Il convient de renforcer les capacités en matière de transfert des connaissances, de renforcement des capacités et d'échange des connaissances dans les domaines où il existe des données probantes suffisantes.
- Il est nécessaire de collecter des données factuelles et d'évaluer les mesures qui produisent des résultats dans les domaines dans lesquels des exemples de solutions existent mais ne sont pas encore étayées par une solide base d'informations.
- Il est nécessaire d'innover en matière d'évaluation de l'apprentissage dans les domaines qui exigent un nouveau type de réflexion et de nouvelles solutions.

Exemples d'activités qui pourraient bénéficier de l'appui du mécanisme KIX

Le tableau 2 présente des exemples d'activités qui pourraient être soutenues par les financements thématiques du mécanisme KIX afin de renforcer les systèmes d'évaluation de l'apprentissage dans les domaines précités. Les lignes pointillées entre les exemples d'activités indiquent qu'un bénéficiaire des financements KIX pourrait éventuellement proposer un projet couvrant de multiples investissements ou possibilités d'avancement.

Tableau 2 : Domaines d'investissements en matière d'échange de connaissances et d'innovations et exemples d'activités

Domaines d'investissements	Opportunités / activités			
		Biens mondiaux à l'appui des institutions nationales d'évaluation des acquis de l'apprentissage	Biens mondiaux à l'appui d'un renforcement de l'expérience en matière d'apprentissage dans l'ensemble du système éducatif	Bien mondiaux à l'appui des systèmes d'évaluation pour les plus marginalisés

Renforcement des capacités nationales	<p>Renforcement des capacités entre les pays pour la conception de cadres harmonisés d'évaluation des apprentissages</p> <p>Échange régional en matière de renforcement institutionnel</p>	<p>Renforcement des capacités des enseignants en matière d'évaluations en classe</p> <p>Renforcement des capacités institutionnelles des responsables des ministères</p>	<p>Participation des pays fragiles ou touchés par un conflit (PFC) aux évaluations internationales</p> <p>Directives pour l'harmonisation des priorités des donateurs en matière d'évaluation</p>	<p>Échanges régionaux et internationaux sur les bonnes pratiques d'utilisation des données d'évaluation de l'apprentissage</p>
Collecte de données probantes et évaluation de ce qui fonctionne	<p>Analyse approfondie des outils de diagnostic des systèmes d'évaluation de l'apprentissage et recommandation des améliorations requises</p> <p>Examen des réformes des examens publics qui donnent de bons résultats</p> <p>Examen des initiatives d'évaluation basée sur les compétences</p>	<p>Examen des bonnes pratiques d'évaluation en classe</p>	<p>Examen de la rentabilité des évaluations</p> <p>Examen de la participation des PFC aux évaluations nationales et internationales</p>	<p>Examen de l'incidence des données d'évaluation de l'apprentissage (évaluations en classe et évaluations à grande échelle) sur les politiques</p>
Innovation	<p>Mise en place conjointe par les groupes de PDP d'évaluations des compétences dépassant la lecture/écriture et le calcul</p>	<p>Innovations peu coûteuses dans le domaine du renforcement durable des compétences des enseignants et des responsables des ministères de l'éducation</p>	<p>Mécanismes d'utilisation des données existantes aux fins d'établissement de rapports au niveau mondial</p> <p>Solutions technologiques peu coûteuses pour la collecte de données dans les PFC, et de données sur les élèves handicapés ou déplacés</p>	<p>Approche « tester-intervenir-tester à nouveau » systématique</p> <p>Nouvelles méthodes de diffusion, de communication et de visualisation des données d'évaluation des acquis de l'apprentissage</p>

Références

- Anderson, Kate, Seamus Hegarty, Martin Henry, Helyn Kim et Esther Care. « *Breadth of Learning Opportunities: A Fresh Approach to Evaluating Education Systems* ». Washington DC: Brookings Institution, 2018.
- Anderson, Kate et Joshua Muskin. « *Learning Champions: How 15 Countries, Cities, and Provinces Came Together to Rethink Learning Assessment* ». Washington DC: Brookings Institution, 2018.
- Archer, Elizabeth. « The Assessment Purpose Triangle: Balancing the Purposes of Educational Assessment ». *Frontiers in Education* 2 (41), 2017.
- Banerji, Rukmini et Michael Walton. « *What Helps Children to Learn? Evaluation of Pratham's Read India Program in Bihar & Uttarakhand* ». New Delhi: Pratham, 2011.
- Banque mondiale. « Student Assessment for Policymakers and Practitioners | World Bank Group ». Consulté le 8 mai 2018. <https://olc.worldbank.org/content/student-assessment-policymakers-and-practitioners>.
- Bernard, Jean-Marc et Katharina Michaelowa. « How Can Countries Use Cross-national Research Results to Address “the Big Policy Issues” ? » (Études de cas en Afrique francophone). In *Cross-national Studies of the Quality of Education: Planning Their Design and Managing Their Impact*, 229-40. Paris: IIEP-UNESCO, 2006.
- Best, Maura, Pat Knight, Petra Leitz, Craig Lockwood, Dita Nugroho et Mollie Tobin. « *The Impact of National and International Assessment Programmes on Education Policy, Particularly Policies regarding Resource Allocation and Teaching and Learning Practices in Developing Countries* ». Melbourne: ACER, 2013.
- Black, Paul et Dylan Wiliam. « Developing a Theory of Formative Assessment ». In J. Gardner, John, éd. *Assessment and Learning*. London: SAGE, 2006.
- Burdett, Newman. « *Review of High Stakes Examination Instruments in Primary and Secondary School in Developing Countries* ». RISE Working Paper 17/018, 2017.
- Care, Esther, Helyn Kim, Kate Anderson et Emily Gustafsson-Wright. « *Skills for a Changing World: National Perspectives and the Global Movement* ». Washington DC: Brookings Institution, 2017.
- Care, Esther, Alvin Vista et Helyn Kim. « *Optimizing Assessment for All* ». Brookings Institution, 2018.
- Chinapah, Vinayagum. « *Monitoring Learning Achievement: Toward Capacity Building* ». Paris: UNESCO, 1997.
- Clarke, Marguerite. « *What Matters Most for Student Assessment Systems: A Framework Paper* ». Washington DC: Banque mondiale, 2012.

- Comité de réflexion sur la métrique des apprentissages. « *Toward Universal Learning: Recommendations from the Learning Metrics Task Force* ». Montréal et Washington DC: ISU et Brookings Institution, 2013.
- . « *Toward Universal Learning: Implementing Assessment to Improve Learning* ». Montréal et Washington DC: ISU et Brookings Institution, 2014.
- Commission internationale pour le financement de possibilités d'éducation. « *La génération d'apprenants : Investir dans l'éducation pour un monde en pleine évolution* ». Washington DC: Commission pour l'éducation, 2016.
- Fischman, Gustavo E., Amelia Marcetti Topper, Iveta Silova, Janna Goebel et Jessica L. Holloway. « An Examination of the Influence of International Large Scale Assessments and Global Learning Metrics on National School Reform Policies ». *Journal of Education Policy*, 7 mai 2018.
- Friedman, Tim, Ursula Schwantner, Jeaniene Spink, Naoko Tabata et Charlotte Waters. « *Improving Quality Education and Children's Learning Outcomes and Effective Practices in the Eastern and Southern Africa Region* ». Nairobi: UNICEF/ Bureau régional de l'Afrique de l'Est et de l'Afrique australe, 2016.
- Greaney, Vincent et Thomas Kellaghan. 2008. « *Assessing National Achievement Levels in Education* ». Washington, DC: Banque mondiale, 2008.
- Groupe de personnalités de haut niveau chargé du programme de développement pour l'après-2015. « *A New Global Partnership: Eradicate Poverty and Transform Economies through Sustainable Development* ». New York : Nations Unies, 2013.
- Groupe thématique 4 du SDN. « *The Future Of Our Children: Lifelong, Multi-Generational Learning For Sustainable Development* ». New York: SDN, 2014.
- IEA. « IEA Studies ». Consulté le 8 mai 2018.
- Institut de statistique de l'UNESCO. « *Indice de la capacité à évaluer l'apprentissage* ». Montréal: Institut de statistique de l'UNESCO. Consulté le 8 mai 2018.
- . « *6 Out of 10 Children and Adolescents Are Not Learning a Minimum in Reading and Math* ». Montréal: Institut de statistique de l'UNESCO, 2017a.
- . « *Investment Case for Expanding Coverage and Comparability for Global Indicator 4.1.1* ». Montréal: Institut de statistique de l'UNESCO, 2017b.
- ISU (Institut de statistique de l'UNESCO) et ACER (Australian Council for Educational Research). « *Principles of Good Practice in Learning Assessment* ». Camberwell, Victoria: ACER, 2017.
- Institut international de l'UNESCO pour le renforcement des capacités en Afrique. « *Teaching policies and learning outcomes in Sub-Saharan Africa: Issues and Options* ». Addis-Abeba: Institut international de l'UNESCO pour le renforcement des capacités en Afrique, 2016.

- Kanjee, Anil. Extraits d'observations issues d'une consultation. 10 août 2018.
- Kelleghan, Thomas, Vincent Greaney et T. Scott Murray. « *Using the Results of a National Assessment of Educational Achievement* ». Vol. 5. Washington, DC : Banque mondiale, 2009.
- Lockheed, Marlaine E. « *The Craft of Education Assessment: Does Participating in International and Regional Assessments Build Assessment Capacity in Developing Countries? An Independent Evaluation of IEA's Program on the Assessment of Student Achievement (PASA)* ». 2010.
- Mugo, John. « *Welcoming Two New Initiatives to Accelerate Quality Teaching and Learning in Africa* ». Washington DC: Brookings Institution, 2017.
- Mutambo, Aggrey. « UK Company Fined Sh330 Million for Bribing Kenya Officials ». *Daily Nation*, 11 janvier 2016.
- OCDE. « *Building Capacity for Assessment in PISA for Development Countries* ». Paris: OCDE, 2017.
- Partenariat mondial pour l'éducation. « *Plan stratégique 2020 du PME* ». Partenariat mondial pour l'éducation. Washington DC: PME, 2016.
- . « *Cadre de résultats du PME pour 2016-2020* ». Partenariat mondial pour l'éducation. Washington DC: PME, 2017.
- Petters, Janet Sunday et Maureen O. Okon. « Students' Perception of Causes and Effects of Examination Malpractice in the Nigerian Educational System: The Way Forward for Quality Education ». *Procedia - Social and Behavioral Sciences* 114: 125-29, 2014.
- Pritchett, Lant. « *Creating Education Systems Coherent for Learning Outcomes: Making the Transition from Schooling to Learning* ». London: Programme RISE, 2014.
- Ramirez, Maria Jose. « *Quick Guide No. 3 Implementing a National Learning Assessment* ». Montréal: ISU, 2017. <http://uis.unesco.org/sites/default/files/documents/quick-guide-3-implementing-national-learning-assessment.pdf>.
- Save the Children. « *Beyond School Walls: A Boost for Readers* ». Fairfield CT: Save the Children, 2013.
- Somerset, H.C.A. « *Examination Reform in Kenya* ». Washington DC: Banque mondiale, 1987.
- Tobin, Mollie, Petra Lietz, Dita Nugroho, Ramya Vivekanandan et Tserennadmid Nyamkhuu. « *Using Large-scale Assessments of Students Learning to Inform Education Policy: Insights from the Asia-Pacific Region* ». Melbourne et Bangkok : ACER et UNESCO Bangkok, 2015.
- Transparency International. « *Global Corruption Report: Education* ». London: Earthscan, Publishing for a Sustainable Future. Routledge, Taylor & Francis Group, 2013.
- UNESCO. « *Rapport mondial de suivi sur l'EPT : Atteindre les marginalisés* ». Paris: UNESCO, 2010.
- . « Policy paper 34: Fulfilling our collective responsibility: Financing global public goods in

education ». Paris: UNESCO, 2018.

USAID. « *Early Grade Reading Barometer* ». Consulté le 8 mai 2018.
<http://www.earlygradereadingbarometer.org/>.

Wagner, Daniel, Sharon Wolf et Robert F. Boruch, éd. « *Learning at the Bottom of the Pyramid: Science, Measurement and Policy in Low-Income Countries* ». Paris: UNESCO IIEP, 2018.

Wagner, Daniel, Martin Lockheed, Ina Mullis, Michael O. Martin, Anil Kanjee, Amber Gove et Amy Dowd. « The Debate on Learning Assessments in Developing Countries ». *Journal of Comparative and International Education*, 42(3): 509-545, 2012.

Winthrop, Rebecca, Eileen McGivney et Adam Barton. « *Can We Leapfrog? The Potential of Education Innovations to Rapidly Accelerate Progress* ». Washington DC: Brookings Institution, 2017.

Annexe A. Fiche d'information du PME sur l'évaluation de l'apprentissage

Aperçu des indicateurs 1, 15 et 20

Trois indicateurs du cadre de résultats sont liés aux évaluations de l'apprentissage :

- Indicateur n° 1 : Proportion de pays en développement partenaires enregistrant une amélioration des acquis scolaires (éducation de base) ;
- Indicateur n° 15 : proportion de PDP dont le système d'évaluation des acquis scolaires dans l'éducation de base répond à des normes de qualité ;
- Indicateur n° 20 : proportion des financements appuyant les systèmes d'information sur la gestion de l'éducation (SIGE) ou les systèmes d'évaluation des apprentissages¹.

Les pays membres du PME sont globalement confrontés à des difficultés ayant trait à la disponibilité des systèmes d'évaluation de l'apprentissage pour suivre les progrès dans le domaine des acquis scolaires (indicateur 1). Trente-deux pour cent seulement des PDP (19 des 60 PDP examinés) étaient dotés de systèmes d'évaluation répondant aux normes de qualité entre 2011 et 2015 (indicateur 15). Pour corriger cette carence, le PME soutient activement le renforcement des systèmes d'évaluation de l'apprentissage dans les PDP par le biais de financements pour la mise en œuvre de programmes sectoriels de l'éducation (ESPIG) (indicateur 20).

Principaux résultats

Le premier jalon de l'indicateur 1 (amélioration des résultats d'apprentissage) est établi pour 2018². L'objectif est fixé à 68 % de l'ensemble des PDP enregistrant une amélioration des acquis scolaires, dont 65 % dans les PFC. Le rapport des résultats de 2016 indiquait que les acquis scolaires s'étaient améliorés dans 13 des 20 pays en développement partenaires du PME pour lesquels des données étaient disponibles à deux moments différents dans le temps (deux PFC sur quatre) pour l'année de référence, entre 2000 et 2015.

Les données de référence de l'indicateur 15 donnent à penser que moins d'un tiers des pays en développement partenaires (32 %) possèdent des systèmes d'évaluation de l'apprentissage (21 % dans le cas des PFC) respectant des normes de qualité³. Le jalon suivant de l'indicateur 15 a été établi pour 2018, avec un objectif de 38 % de l'ensemble des PDP possédant un système d'évaluation de l'apprentissage respectant des normes de qualité, dont 29 % des PFC. Il n'existe pas de nouvelles données pour cet indicateur pour 2017.

¹ Il n'y a pas de ventilation par sexe pour ces trois indicateurs du cadre de résultats.

² La fiche méthodologique de l'indicateur 1 est disponible (en anglais) à l'adresse suivante : <https://www.globalpartnership.org/content/methodology-sheet-PME-result-indicator-1>.

³ La fiche méthodologique de l'indicateur 15, qui s'inspire de l'« Approche systémique pour de meilleurs résultats dans le domaine de l'éducation (SABER) » de la Banque mondiale, mais qui a été contextualisée en fonction des besoins des PDP du PME, est disponible (en anglais) à l'adresse suivante : <https://www.globalpartnership.org/content/methodology-sheet-PME-result-indicator-15>.

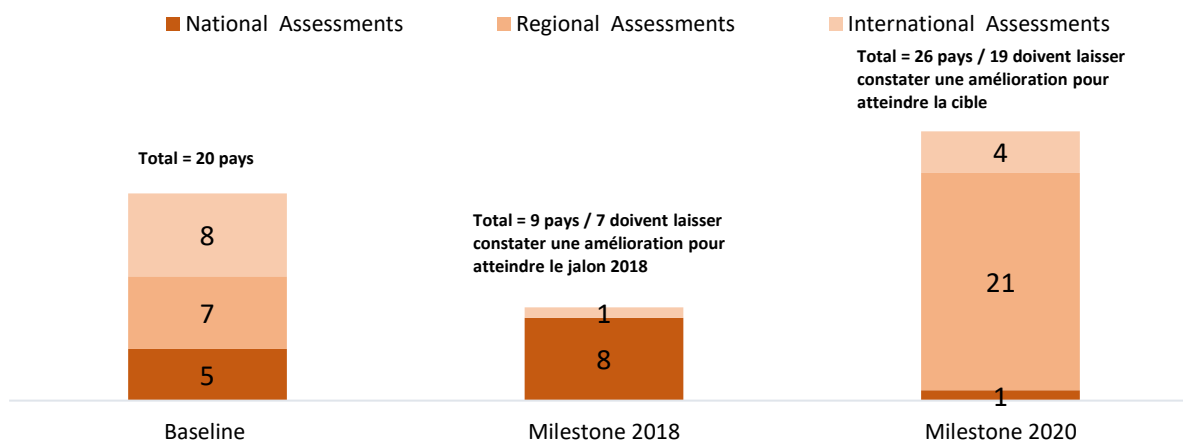
Aucun jalon n'a été établi pour la proportion des ESPIG soutenant un SIGE ou un autre système d'évaluation de l'apprentissage (indicateur 20) pour 2017⁴. Toutefois, les données donnent à conclure que 92 % (44 sur 48) des ESPIG actifs au cours de l'exercice 17 soutenaient un SIGE ou un autre système d'évaluation de l'apprentissage, dépassant ainsi de loin le premier jalon de l'indicateur fixé à 50 % pour 2018⁵. Dans le cas des PFC, cette proportion s'établit à 96 % (26 sur 27), dépassant encore une fois et de loin le jalon de l'indicateur fixé à 44 % pour 2018.

Évaluations des progrès en matière d'acquis scolaires

Bien qu'aucun jalon n'ait été fixé pour l'indicateur 1 en 2017, le PME contrôle la disponibilité des données sur l'évaluation de l'apprentissage. Plusieurs évaluations internationales, régionales et nationales auront été réalisées entre 2011 et 2019 dans les PDP, ce qui permettra au PME de mesurer les progrès en matière d'acquis scolaires et d'en faire état dans des rapports. Le PME estime ainsi que neuf pays auront participé deux fois aux mêmes évaluations entre 2011 et 2017, et que 26 pays auront fait de même entre 2011 et 2019 (figure 1).

Figure 1 : Le volume des données permettant d'évaluer les acquis de l'apprentissage augmentera légèrement d'ici 2020

Nombre de pays disposant de données d'évaluation des acquis de l'apprentissage en 2018 et 2020



Le PME renforce son appui en faveur de l'évaluation de l'apprentissage, comme le laisse constater la hausse de la proportion des ESPIG soutenant des systèmes d'évaluation de l'apprentissage ou des SIGE. Tel que mentionné ci-dessus, les données actuelles sur l'indicateur 20 donnent à conclure que 92 % (44 sur 48) des ESPIG actifs⁶ au cours de l'exercice 17 soutenaient un SIGE ou un autre système d'évaluation de l'apprentissage, contre 83 % (45 sur 54) au cours de l'exercice 16 (figure 2)⁷.

⁴ La fiche méthodologique de l'indicateur 20 est disponible (en anglais) à l'adresse suivante : <https://www.globalpartnership.org/content/methodology-sheet-PME-result-indicator-20>.

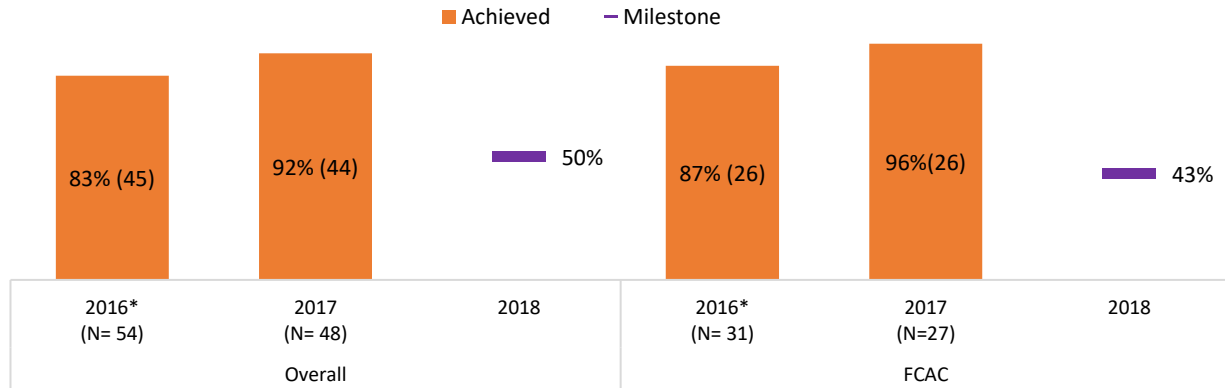
⁵ Incluant sept ESPIG financés par un fonds commun.

⁶ Incluant sept ESPIG financés par un fonds commun.

⁷ La méthode de notation de l'indicateur 20 a changé pour l'exercice 17. Les données correspondant à l'exercice 17 ont été obtenues par recodage des données de l'exercice 16 en utilisant la méthode de notation retenue pour l'exercice 17.

Figure 2 : Presque tous les financements du PME soutiennent un SIGE ou un autre système d'évaluation de l'apprentissage

Proportion des financements actifs soutenant un SIGE ou un autre système d'évaluation de l'apprentissage, exercices 16 et 17



Note : Les données de l'exercice 16 sont obtenues par recodage.

Annexe B. Évaluations nationales et internationales réalisées dans les PDP, de 2010 à 2017

Pays	Types d'évaluation des acquis réalisés par année				
	Nationale	Régionale	Internationale	Évaluation citoyenne	Petites classes
Afghanistan	2013				EGRA 2015
Albanie	2015		PISA 2012, 2015		
Bangladesh	2013			2015	
Bénin	2011	PASEC 2014			EGRA 2015
Bhoutan	2013				
Burkina Faso	2014	PASEC 2014			
Burundi	2011	PASEC 2014			EGRA 2011
Cambodge	2011	PASEC 2014 *SEA-PLM 2016			EGRA 2016
Cameroun	2011	PASEC 2014			EGRA 2015
République centrafricaine					
Tchad		PASEC 2014			EGRA 2012
Comores					
République démocratique du Congo					EGRA 2015, EGMA 2015
République du Congo		PASEC 2014			
Côte d'Ivoire	2011	PASEC 2014			
Djibouti					
Dominique					
Érythrée	2015				
Éthiopie	2013				EGRA 2014
Gambie	2014				EGRA 2013
Géorgie	2015		PIRLS 2010, 2015 TIMSS 2010, 2015		
Ghana	2014		TIMSS 2010	2016	EGRA 2015, EGMA 2013
Grenade					
Guinée	2014				
Guinée-Bissau	2014				
Guyane	2014				
Haïti			*LaNA 2016		EGRA 2015
Honduras	2014	TERCE 2013	PIRLS 2010 TIMSS 2010		EGRA 2016

Pays	Types d'évaluation des acquis réalisés par année				
	Nationale	Régionale	Internationale	Évaluation citoyenne	Petites classes
Kenya		SEACMEQ 2013		2015	EGRA 2015 EGMA 2012
République kirghize			PISA 2015		EGRA 2011
RDP lao	2012	PASEC 2014 *SEA-PLM 2016			EGRA 2016
Lesotho	2014	SEACMEQ 2013			EGRA 2015
Libéria					EGRA 2015 EGMA 2015
Madagascar	2012				
Malawi	2012	SEACMEQ 2013			EGRA 2016 EGMA 2011
Mali		PASEC 2014		2016	EGRA 2015, EGMA 2011
Mauritanie					
Moldavie			PISA 2015		
Mongolie	2013				
Mozambique	2013	SEACMEQ 2013		2016	EGRA 2015
Népal	2013				EGRA 2014
Nicaragua	2010	TERCE 2013			EGRA 2015 EGMA 2011
Niger		PASEC 2014			EGRA 2014
Nigéria	2011				EGRA 2015 EGMA 2013
Pakistan	2014			2016	EGRA 2014
Papouasie-Nouvelle-Guinée		PILNA 2012			EGRA 2013
Rwanda	2011				EGRA 2015 EGMA 2011
Sainte-Lucie					
Saint-Vincent-et-les Grenadines					
São Tomé-et-Principe	2015				
Sénégal	2014	PASEC 2014		2013	EGRA 2014
Sierra Leone			MICS 2018		EGRA 2012 EGMA 2014
Somalie					EGRA 2013
Soudan du Sud					EGRA 2013

Pays	Types d'évaluation des acquis réalisés par année				
	Nationale	Régionale	Internationale	Évaluation citoyenne	Petites classes
Soudan	2015				EGRA 2015
Tadjikistan					EGRA 2016
Tanzanie	2014	SEACMEQ 2013		2012	EGRA 2016 EGMA 2016
Timor-Leste					EGRA 2011
Togo	2013	PASEC 2014	MICS 2018		EGRA 2015
Ouganda	2013	SEACMEQ 2013		2015	EGRA 2015
Ouzbékistan					
Viet Nam	2013	PASEC *SEA-PLM 2016	PISA 2012, 2015		EGRA 2010
Yémen			TIMSS 2010		
Zambie	2013	SEACMEQ 2013			EGRA 2016 EGMA 2011
Zimbabwe	2014	SEACMEQ 2013			EGRA 2013
Total (65)	39	25	11	11	41

L'astérisque (*) indique qu'il s'agit uniquement d'un essai de terrain. EGRA = évaluations de l'apprentissage précoce de la lecture ; EGMA = évaluation de l'apprentissage précoce des mathématiques ; LaNA = évaluation des aptitudes à la lecture, à l'écriture et au calcul ; MICS = Enquête par grappes à indicateurs multiples : module sur les compétences fondamentales d'apprentissage ; PASEC = Programme d'analyse des systèmes éducatifs de la CONFEMEN ; PILNA = évaluation régionale des aptitudes à la lecture, à l'écriture et au calcul de la communauté du Pacifique ; PIRLS = Programme international de recherche en lecture scolaire ; PISA = Programme international pour le suivi des acquis des élèves ; SEA-PLM = initiative pour la mesure de l'apprentissage à l'école primaire en Asie du Sud-est ; SEACMEQ = Consortium d'Afrique australe pour le suivi de la qualité de l'éducation ; TERCE = Troisième étude explicative et comparative régionale ; TIMSS = Étude internationale sur les tendances de l'enseignement des sciences et des mathématiques.

Note : Les PDP qui n'ont pas réalisé les évaluations énumérées sont indiqués en orange dans le tableau ci-dessus.

Source : Catalogue de l'ISU sur les mesures des acquis de l'apprentissage (CLA)/Indice de la capacité à évaluer l'apprentissage (LACI ; Réseau PAL, dispositifs de suivi EGRA et EGMA (RTI International)).