

1 UIS/EDS/7

DONNÉES DES ÉVALUATIONS D'APPRENTISSAGE ET ENQUÊTES SUR LES COMPÉTENCES

DÉFIS ET SOLUTIONS PROPOSÉES

FÉVRIER 2024

2024 CONFÉRENCE SUR LES
DONNÉES ET STATISTIQUES
DE L'ÉDUCATION

INTRODUCTION

Il existe plusieurs décennies de développement et de débats sur la mesure des progrès en matière d'acquis d'apprentissage et de compétences. Ce document analyse et évalue l'état et les lacunes des cadres et méthodologies qui aident à mesurer et à suivre les indicateurs sur les résultats d'apprentissage et les compétences liés aux cibles des ODD 4.1, 4.4, 4.6 et 4.7 (Tableau 1). Il vise à identifier les domaines qui nécessitent une attention et des améliorations supplémentaires. Plus spécifiquement, le document vise à apporter une contribution significative aux discussions et initiatives en cours dans ce domaine, visant à terme à établir une communauté de pratique internationale capable de relever collectivement les défis à venir.

Tableau 1. Indicateurs de l'ODD 4 liés aux acquis d'apprentissage et aux compétences

Indicateur		Domaine	Définitions requises
4.1.1	Proportion d'enfants et de jeunes a) en 2e ou 3e année de cycle primaire ; b) en fin de cycle primaire ; c) en fin de premier cycle du secondaire qui maîtrisent au moins les normes d'aptitudes minimales en i) lecture et ii) mathématiques, par sexe	Lecture et mathématiques	Niveau de compétence minimum Qualité procédurale minimale
4.2.1	Proportion d'enfants âgés de 24 à 59 mois dont le développement est en bonne voie en matière de santé, d'apprentissage et de bien-être psychosocial, par sexe	Apprentissage, santé socio-émotionnelle	Développement en bonne voie
4.4.2	Pourcentage de jeunes/adultes ayant atteint au moins un niveau minimum de compétence en littéracie numérique	Compétences en littéracie numérique	Compétences numériques pertinentes pour l'emploi, les emplois décents et l'entrepreneuriat
4.6.1	Proportion de la population d'un groupe d'âge donné ayant les compétences voulues à au moins un niveau d'aptitude fixé a) en alphabétisme et b) arithmétique fonctionnels, par sexe	Alphabétisation et calcul	Niveau fixé en alphabétisme et arithmétique fonctionnels
4.7.4	Pourcentage d'élèves du premier cycle du secondaire montrant une compréhension adéquate des questions liées à la citoyenneté mondiale et à la durabilité	Citoyenneté mondiale et durabilité	Compréhension adéquate de la citoyenneté mondiale et de la durabilité
4.7.5	Pourcentage d'élèves du premier cycle du secondaire démontrant des compétences en sciences de l'environnement et les géosciences	Sciences de l'environnement et géosciences	Maîtrise des connaissances en sciences de l'environnement et en géosciences

ÉVALUATION DE L'ÉTAT DE MESURE DE L'INDICATEUR ODD 4.1.1

L'indicateur ODD 4.1.1 est la proportion d'enfants et de jeunes (a) en 2^e ou 3^e année de cycle primaire ; (b) en fin de cycle primaire ; et (c) en fin de premier cycle du secondaire, qui maîtrisent au moins le niveau de compétences minimum (MPL) en (i) lecture et (ii) mathématiques, par sexe. Il fait référence à trois niveaux d'enseignement – premières années du primaire (2/3), fin du primaire et premier cycle du secondaire – et à deux matières – la lecture et les mathématiques.

Le format de 'rapport' de l'indicateur vise à communiquer deux informations :

- Le pourcentage d'élèves atteignant au moins un niveau minimum de compétence (MPL) pour les domaines concernés (mathématiques et lecture) pour chaque point de mesure (années 2/3 ; fin du primaire et fin du premier cycle du secondaire).
- Si un programme peut être considéré comme comparable, et les conditions dans lesquelles le pourcentage d'enfants atteignant ou dépassant le niveau de compétence minimum peut être considéré comme comparable au pourcentage déclaré dans un autre pays.

L'indicateur nécessite les entrées suivantes :

- Domaines : Lecture et mathématiques.
- Niveau de compétence minimum (MPL) : référence de connaissances de base dans les domaines à un âge/niveau donné.
- Échantillon : représentatif de la population concernée.
- Procédures : Conformes aux normes minimales de qualité.

Défis

Il existe quelques problèmes critiques dans le rapport de l'indicateur 4.1.1.

Comparabilité des années d'études et des niveaux d'éducation

Étant donné que la durée de la scolarité primaire varie selon les pays, des termes tels que « fin du primaire » peuvent avoir des significations différentes selon les endroits, ce qui complique les

comparaisons entre les pays et les programmes d'évaluation. Cependant, 89 % des pays terminent leur cycle primaire entre la 5e et la 7e année, le problème pourrait donc être mineur.

Comparabilité des résultats d'évaluation dans l'espace et dans le temps

Même si la comparabilité des statistiques entre pays influence la comparabilité dans le temps, cette dernière n'implique pas la première.

- La comparaison entre pays au moyen d'évaluations transnationales facilite la comparabilité entre les pays, à un moment donné. Si chaque programme d'évaluation produit des statistiques comparables dans le temps, alors les statistiques seraient comparables dans le temps et pays.
- Les programmes nationaux d'évaluation ne sont pas comparables les uns aux autres de par leur conception, mais ils peuvent néanmoins fournir des données sur les tendances relativement fiables si la qualité des mesures est suffisamment bonne.

Actualité et impact politique des statistiques

Les évaluations produisent des statistiques nationales, et souvent infranationales, qui peuvent influencer positivement l'élaboration et la mise en œuvre des politiques. Pour que ces impacts positifs se fassent sentir, les statistiques doivent non seulement être exactes, mais également être largement considérées comme crédibles, et le délai d'exécution entre l'évaluation et la communication des résultats doit être aussi court que possible.

Qualité procédurale

Des opérations et des procédures robustes et cohérentes constituent un élément essentiel de toute évaluation à grande échelle, afin de maximiser la qualité des données et de minimiser l'impact des variations procédurales sur les résultats. Des exemples de normes procédurales peuvent être trouvés dans toutes les évaluations internationales à grande échelle et dans de nombreuses évaluations au niveau régional à grande échelle, où l'objectif est d'établir une cohérence procédurale entre les pays. De nombreuses évaluations nationales établissent également des directives procédurales claires pour garantir la cohérence de leur mise en œuvre.

La mise en œuvre de l'évaluation est confrontée à de nombreuses décisions méthodologiques, notamment les formats de test et les décisions d'échantillonnage. Il n'est pas nécessaire de

recourir à des procédures et des formats identiques pour toutes les évaluations. Cependant, un ensemble minimum de procédures est nécessaire – l'alignement des procédures – afin que l'intégrité des données soit protégée et que les résultats soient robustes et raisonnablement comparables pour un pays donné au fil du temps, ainsi qu'entre les pays à un moment donné.

Coûts financiers des évaluations pour les pays

Les évaluations sont relativement coûteuses. Cependant, même pour les pays en développement, le coût d'évaluer les résultats systématiquement est extrêmement faible relativement au coût global de fournir la scolarité sans mesurer les résultats d'apprentissage et les compétences.¹

Faible couverture des évaluations transnationales dans les pays à revenu faible et intermédiaire inférieur

L'indicateur ODD 4.1.1 est rapporté à l'aide de diverses études transnationales internationales ([PIRLS](#), [TIMSS](#)) ou régionales ([PILNA](#), [SEA-PLM](#), [PASEC](#), [LLECE](#), [SACMEQ](#)) (**Tableau 2**). Ces outils n'ont pas été conçus pour le rapport des ODD mais, en 2018, l'Alliance mondiale pour le suivi de l'apprentissage ([GAML](#)) et le Groupe de coopération technique sur les indicateurs de l'ODD 4 ([TCG](#)) ont convenu que ces évaluations pourraient être utilisées pour rendre compte de l'apprentissage en fonction de leurs niveaux de compétence qui « correspondent » le mieux au MPL global.

¹Pour plus d'informations sur le calcul des coûts, voir ISU (2023). *Reporting learning outcomes in basic education: Country's options for indicator 4.1.1* : https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option_Zambia_2023.05.15_FINAL.pdf.

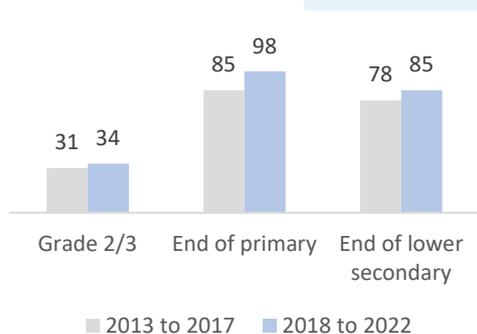
Tableau 2. Programmes d'évaluation par année d'études ou âge et utilisation pour le rapport sur l'indicateur ODD 4.1.1

Année d'études/âge	Programme d'évaluation international
ODD 4.1.1a : Premières années	
2	PASEC
3	ERCE , AMPLa
ODD 4.1.1b : Fin du primaire	
4	PILNA , LaNA , PIRLS , TIMSS
5	SEA-PLM
6	LaNA , PASEC , PILNA , SACMEQ , ERCE , AMPLb
ODD 4.1.1c : Fin du premier cycle du secondaire	
8	TIMSS
Âge : 15 ans	PISE

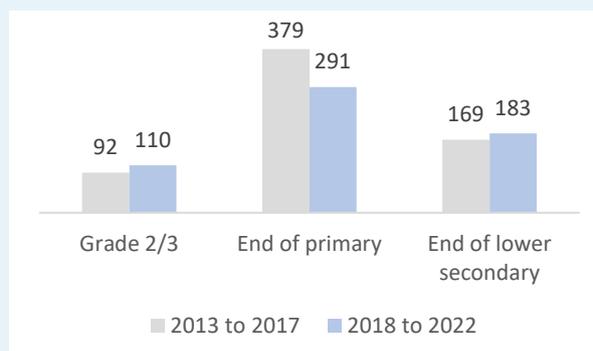
Toutefois, la production de résultats d'apprentissage comparables ne progresse pas suffisamment rapidement et de manière égale. Quel que soit le critère de couverture (le nombre de pays ou la population), la couverture est beaucoup plus élevée à la fin du primaire et à la fin du premier cycle du secondaire qu'en 2e ou 3e année (**Figure 1**).

Figure 1. Couverture des évaluations des apprentissages, par niveau d'éducation

à. Nombre de pays



b. Population d'âge scolaire en millions



Évaluations nationales : alignement et qualité des procédures

Bien que les données de nombreuses évaluations nationales des apprentissages soient facilement disponibles, chaque pays fixe ses propres normes, ce qui conduit à des définitions incohérentes des niveaux de performance. L'analyse des résultats reste donc limitée à un seul test, une seule méthodologie et une seule échelle. Même si les méthodologies tendent à

converger entre les évaluations internationales et régionales, il est encore difficile d'établir un niveau de référence commun pour les évaluations nationales et l'ensemble des procédures utilisées pour l'échantillonnage, la gestion des données et les rapports diffèrent également.

Le deuxième aspect auquel les programmes nationaux d'évaluation doivent se conformer est la qualité des procédures. Les informations sur les processus des programmes nationaux doivent être accessibles au public et suffisamment détaillées. L'ISU a fourni des conseils et des outils d'auto-évaluation concernant cet aspect dans [l'Alignement et le reporting sur l'indicateur 4.1.1 : Flux de travail annoté de l'ISU](#).

Normes

Compte tenu de ces défis, l'ISU s'est concentré sur la définition (i) du niveau de compétence minimum, en l'alignant sur un concept de compétence indépendant d'un cadre d'évaluation particulier, d'éléments ou de tests spécifiques, pour permettre l'établissement de rapports et (ii) d'un ensemble de stratégies de liaison au cadre de compétences.

Niveau de compétence minimum

Le niveau de compétence minimum (MPL) est la référence des connaissances de base dans un domaine (par exemple, mathématiques, lecture) à un âge/niveau donnés, mesurée au moyen d'évaluations de l'apprentissage. Le MPL est un point de référence pour rendre compte des compétences minimales à chaque niveau scolaire, sans nécessiter un seul test pour résoudre la comparabilité.

Le descripteur de niveau de compétence MPL est la norme clé pour chaque niveau et domaine qui permet l'utilisation de plusieurs évaluations pour rendre compte de l'indicateur² (**Tableau 3**).

²Il a été convenu de rendre compte conformément à la définition textuelle du niveau de compétence minimum pour chaque domaine et niveau dans les évaluations transnationales. Ceci a été établi en effectuant une analyse des descripteurs de niveau de performance de ces évaluations en lecture et en mathématiques.

En 2018, un accord a été conclu avec les programmes d'évaluation transnationaux sur lequel de leurs descripteurs était le mieux aligné sur le descripteur MPL, bien qu'une validation plus approfondie soit nécessaire, par le biais d'un exercice d'établissement de normes, car les programmes d'évaluation n'ont pas été conçus pour mesurer ODD 4.

Tableau 3. Niveaux de compétence minimaux en lecture et en mathématiques pour l'indicateur 4.1.1

Niveau d'éducation	Descripteur	
	Lecture	Mathématiques
2e année	Les élèves lisent et comprennent la plupart des mots écrits, en particulier ceux qui sont familiers, et extraient des informations explicites des phrases.	Les élèves démontrent des compétences en sens du nombre et en calcul, en reconnaissance de formes et en orientation spatiale.
3e année	Les élèves lisent à haute voix des mots écrits avec précision et fluidité. Ils comprennent le sens global des phrases et des textes courts.	
4e à 6e année	Les élèves interprètent et donnent des explications sur les idées principales et secondaires dans différents types de textes. Ils établissent des liens entre les idées principales d'un texte et leurs expériences personnelles ainsi que leurs connaissances générales.	Les élèves démontrent des compétences en sens du nombre et en calcul, en mesure de base, en lecture, en interprétation et en construction de graphiques, en orientation spatiale et en modèles numériques.
8e à 9e années	Les élèves établissent des liens entre les idées principales de différents types de texte et les intentions de l'auteur. Ils réfléchissent et tirent des conclusions à partir du texte.	Les étudiants démontrent des compétences en calcul, en problèmes d'application, en faisant correspondre des tableaux et des graphiques et en utilisant des représentations algébriques.

Source : Institut de statistique de l'UNESCO, 2021, p. 4.

Dans le cadre de cette démarche, il a fallu :

- Estimer approximativement les niveaux d'intérêt pour le rapport
- Utiliser les niveaux de compétence existants du programme d'évaluation
- Identifier le descripteur de niveau de compétence qui correspond le mieux aux niveaux de compétence minimaux
- Utiliser ce niveau pour faire rapport jusqu'à ce que l'exercice d'établissement de normes soit finalisé

- Renforcer les capacités techniques nationales, mais ne pas aborder directement le développement des évaluations nationales.

Cadre mondial de compétence

Le cadre mondial de compétence (GPF) est une référence mondiale utile, qui définit une référence commune pour les niveaux de compétence en lecture et en mathématiques que les apprenants sont censés démontrer à la fin de chaque niveau scolaire, de la 1^{re} à la 9^e année. Les quatre niveaux décrits dans le GPF – « en dessous du niveau de compétence minimal mondial », « atteint partiellement le niveau de compétence minimal mondial », « atteint le niveau de compétence minimal mondial » et « dépasse le niveau de compétence minimal mondial » - forme une échelle commune allant de réalisation faible à élevée. Le GPF et ses niveaux de compétence associés donnent des indications sur l'ensemble minimum de compétences que les élèves doivent acquérir sur le chemin menant à la maîtrise de la lecture et des mathématiques.

Lier les programmes d'évaluation au niveau de compétence minimum

Lier une évaluation nationale, régionale ou internationale à la définition mondiale du MPL nécessite une méthodologie permettant d'identifier les mêmes concepts et définitions dans des programmes d'évaluation conçus à des fins complètement différentes afin de permettre un certain degré de comparabilité et des déductions justes sur les pays comparés.

Le processus de comparaison de différentes évaluations (modération) peut être statistique ou non statistique. Trois processus sont le projet Rosetta Stone, la méthodologie de liaison des politiques et les évaluations des niveaux minimum de compétence (AMPL).³

³Pour plus d'informations sur les stratégies liées, leurs coûts, leurs avantages, l'état d'exécution, les jalons exécutés et en attente et les délais, voir UIS (2023) *Reporting learning outcomes in basic education: Country's options for indicator 4.1.1*

Pierre de Rosette

Le projet Pierre de Rosette (Rosetta Stone), dirigé par l'Association internationale pour l'évaluation des résultats scolaires (IEA), a tenté d'harmoniser les données de différentes évaluations. Nommé d'après la célèbre découverte archéologique qui a permis la traduction entre différentes langues écrites, il est conçu pour relier les évaluations régionales aux évaluations internationales de l'apprentissage. L'objectif est de fournir aux pays qui ont participé à des évaluations régionales (ou nationales), mais pas à des évaluations internationales, des informations sur la proportion d'élèves ayant obtenu le MPL en lecture et en mathématiques. Dans un premier effort pour mettre en œuvre cette approche et établir des tables de concordance, les évaluations régionales Etude Comparative et Explicative Régionale (ERCE) et Programme d'analyse des systèmes éducatifs de la CONFEMEN ([PASEC](#)) ont été liés aux Trends in International Mathematics and Science Study ([TIMSS](#)) pour les mathématiques et au Progress in International Reading Literacy Study ([PIRLS](#)) pour la lecture.

Liaison de politiques (Policy Linking)

Une autre approche pour harmoniser les évaluations est la méthodologie de liaison de politiques, qui est une méthode non statistique qui utilise le jugement pour aligner et faire correspondre les éléments de l'évaluation nationale avec le GPF. Ce processus établit des références mondiales comparables au niveau international sur la base des descripteurs de chaque référence GPF. Trois tâches principales – alignement, correspondance et définition de références – doivent être réalisées au cours d'un atelier de 5 à 6 jours avec 15 à 20 panélistes (enseignants), experts en programmes d'études et en évaluation de chaque niveau/matière, pour identifier et définir, si possible, les références requises pour les rapports internationaux sur l'indicateur ODD 4.1.1.

Afin de produire des références fiables pour les rapports internationaux, le Policy Linking Toolkit spécifie cinq critères, notamment qu'un nombre suffisant d'éléments nationaux soient alignés

gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option_Zambia_2023.05.15_FINAL.pdf.

sur le GPF ; les échantillons sont représentatifs à l'échelle nationale ; et les évaluations nationales sont administrées selon des normes de qualité minimales. La méthodologie de liaison de politiques a été proposée en 2017, testée en 2019, révisée en 2020 et testée à nouveau en 2021-2022. La Toolkit a ensuite été révisée en 2023 et est désormais en phase pilote.

Évaluations des niveaux de compétence minimaux

Des évaluations des niveaux de compétence minimaux (AMPL) ont également été proposées comme moyen d'obtenir des données pouvant être utilisées pour mesurer et suivre l'indicateur ODD 4.1.1. Les AMPL sont des outils robustes destinés à mesurer l'atteinte du MPL en lecture et en mathématiques à un niveau donné du cycle éducatif. Cela permet aux pays de produire des données sur les résultats d'apprentissage comparables au niveau international pour rendre compte de l'indicateur mondial 4.1.1 des ODD.

L'AMPL-a, qui mesure les compétences dans les premières années d'études, est en cours de développement et sera testé et administré en 2023, en anglais et en français. L'AMPL-b, qui mesure les compétences à la fin du primaire, a été développé en 2021 en anglais et en français et a été mis en œuvre dans six pays africains dans le cadre du projet *Monitoring Impacts on Learning Outcomes* : Burkina Faso, Burundi, Côte d'Ivoire, Kenya, Sénégal et Zambie. AMPL-b a été administré en tant que module autonome en Sierra Leone en 2022 et sa mise en œuvre est prévue en Jordanie et au Pakistan.

Alternatives des pays en matière de rapport

Pour guider le choix de la mesure de l'apprentissage et garantir que les données d'évaluation sont cohérentes avec les objectifs stratégiques à long terme d'une prise de décision efficace, l'ISU, l'UNESCO, la Banque mondiale et l'UNICEF ont élaboré un ensemble de principes - le Learning Data Compact. Ces principes sont importants non seulement pour concevoir des évaluations ou décider quelle évaluation utiliser « prête à l'emploi », mais également pour développer des systèmes d'évaluation nationaux. Ces principes s'appuient sur ceux qui existent déjà ; permettre la flexibilité nécessaire pour garantir l'alignement sur les besoins du pays (il ne s'agit pas d'une solution universelle) ; favoriser l'appropriation nationale grâce à une approche axée sur la demande ; et garantir que les données sont pertinentes pour la prise de décision.

Un système d'évaluation national devrait être capable de produire des rapports, de gérer l'amélioration à tous les niveaux de l'éducation, de guider la prise de décision et de relier les évaluations au niveau du système aux évaluations formatives et aux pratiques en classe. Pour garantir que les évaluations puissent suivre avec précision les progrès en vue de la prise de décision, les données doivent être comparables au niveau international. Chaque pays devrait disposer d'une évaluation conçue ou pouvant être utilisée à des fins de comparabilité internationale – un engagement dans le processus des ODD. Les options de rapport des pays sont diverses (**Tableau 4**), mais le choix doit être guidé par la ou les évaluations adaptées à l'objectif et les plus rentables, en tenant compte de la situation initiale du pays et de l'objectif d'avoir une comparabilité dans le temps et la représentativité des résultats au niveau national.

Deux cas particuliers sont à noter. Les évaluations nationales pourraient être utilisées pour établir des rapports, sous réserve de l'utilisation de liens statistiques qui pourraient être mis en œuvre à l'aide de modules calibrés tels que AMPL. Des outils tels que le MPL et le GPF servent à comprendre et à comparer les normes mondiales, tandis que la liaison de politiques peut inciter les parties prenantes nationales à analyser l'évaluation par rapport à ces normes. L'ISU prend en compte les pays qui souhaitent rendre compte à l'échelle mondiale sur l'indicateur 4.1.1. L'ISU ajoute un module calibré à l'évaluation nationale, tel que l'AMPL, et le complète par la liaison de politiques à des fins de renforcement des capacités, étant donné que la méthodologie est encore dans une phase pilote.

Un deuxième cas concerne les premières années d'études, ou 4.1.1a, où des outils tels que l'évaluation précoce de la lecture et des mathématiques (EGRA/EGMA), les évaluations citoyennes du réseau PAL et le module d'apprentissage fondamental de l'UNICEF de son enquête MICS auprès des ménages pourraient être pertinents et utilisés pour le rapport. Même si ces évaluations ont été appliquées à l'échelle mondiale, elles ne peuvent actuellement pas être utilisées pour des rapports mondiaux, principalement parce qu'elles n'étaient pas destinées à générer des données comparables. Cependant, elles ont le potentiel d'être utilisées pour des rapports mondiaux et l'ISU étudie la manière de tirer le meilleur parti de ces évaluations.⁴

Tableau 4. Alternatives aux pays pour le rapport sur l'indicateur ODD 4.1.1

	4.1.1.a	4.1.1.b	4.1.1.c	Couverture
Évaluations nationales – liens statistiques via des modules calibrés				
AMPL	•	•		
Module PISA			•	
Participation aux évaluations transnationales				
PILNA		•		îles du Pacifique
PASEC	•	•	•	Principalement Afrique (francophone)
SACMEQ		•		Afrique (sud et est)
SEA-PLM		•		Asie du sud-est
LLECE	•	•		l'Amérique latine
TIMSS	•	•	•	Mondial
PIRLS	•	•		Mondial
PISE			•	Mondial

⁴Voir <https://world-education-blog.org/2023/09/13/compare-align-track-the-foundational-learning-data-challenge/>.

ÉVALUATION DE L'ÉTAT DE MESURE DES INDICATEURS 4.6.1, 4.7.4 ET 4.7.5

Des progrès significatifs ont été réalisés dans la mise en place de cadres conceptuels, méthodologiques et de rapport pour les indicateurs 4.6.1, 4.7.4 et 4.7.5. Les cadres fournissent une approche structurée et proposent des lignes directrices et des principes pour la collecte et l'analyse des données. Cependant, malgré les progrès réalisés dans la définition de cadres depuis 2015, la faible couverture des données constitue un problème majeur. Relever ce défi nécessite un effort concerté pour établir des définitions et des mesures communes, garantissant une approche standardisée dans toutes les évaluations (**Tableau 5**).

Tableau 5. Disponibilité d'évaluations pour mesurer les indicateurs des ODD 4.6.1, 4.7.4 et 4.7.5

Indicateurs	Métadonnées	Sources de données	Dernière administration	Durée du cycle	Couverture de pays (N)	Population couverte (%)
4.6.1	Oui	Programme pour l'évaluation internationale des compétences des adultes (PIAAC)	2017	10 années	37	21
4.7.4	Oui	Étude internationale sur l'éducation civique et la citoyenneté (ICCS)	2016	6/7 ans	23	10
4.7.5	Oui	TIMSS, PISA	2019/2022	4/5 ans	38	16

Les pays à revenu faible ou intermédiaire manquent souvent des ressources et des infrastructures nécessaires pour développer et mettre en œuvre des outils permettant de mesurer efficacement ces indicateurs. Pour résoudre ce problème, il faut non seulement développer des outils de mesure appropriés, mais également des initiatives ciblées de renforcement des capacités dans les pays à revenu faible et intermédiaire pour garantir que ces indicateurs sont mesurés de manière complète et précise.

L'ISU a proposé une solution pour accroître le rapport de l'indicateur mondial 4.6.1 des ODD. S'appuyant sur son expérience avec le Programme d'évaluation et de suivi de l'alphabétisation (LAMP), il a développé une version réduite de l'enquête afin de réduire la charge opérationnelle, financière et technique. En particulier, cette mini-enquête LAMP permettrait, entre autres ajustements, de réduire le nombre de domaines de compétences évalués, d'adapter les procédures d'échantillonnage et d'offrir la possibilité d'introduire l'enquête comme un module dans une enquête auprès des ménages existante, plutôt que comme une enquête autonome.⁵

Il n'existe actuellement aucune alternative pour accroître le rapport sur les indicateurs des ODD 4.7.4 et 4.7.5, si ce n'est qu'un plus grand nombre de pays se joignent aux évaluations transnationales qui les mesurent. Cependant, une approche similaire d'un module axé sur la mesure de la norme mondiale convenue concernant ce qu'est une compréhension adéquate des questions liées à la citoyenneté mondiale et à la durabilité (indicateur 4.7.4) et, de la même manière, la maîtrise des connaissances en sciences de l'environnement et en géosciences (indicateur 4.7.5) pourrait être explorée.

QUESTIONNAIRES CONTEXTUELS

Au-delà de l'harmonisation des données sur les résultats d'apprentissage et les compétences, l'harmonisation des questionnaires contextuels issus de différentes évaluations à grande échelle est d'une importance cruciale pour permettre des analyses comparatives solides des tendances, des modèles et des déterminants des inégalités éducatives entre les pays et au fil du temps.

Si l'objectif principal des évaluations transnationales est d'estimer les résultats d'apprentissage, elles collectent également un riche ensemble d'informations de base :

- Elles recueillent auprès des **élèves** des informations sur leurs expériences scolaires, leurs attitudes à l'égard des matières enseignées et les caractéristiques de leurs parents et de leur

⁵Voir <https://uis.unesco.org/sites/default/files/documents/mini-lamp-monitoring-progress-sdg4.6.1-2018-en.pdf>.

foyer, en plus des informations démographiques de base, telles que l'âge et le sexe (**Tableau 6**).

- Elles recueillent auprès des **enseignants** des informations sur leurs attitudes à l'égard de l'enseignement, leurs opinions sur les ressources pédagogiques, leur formation scolaire et leur développement professionnel continu.
- Elles collectent des **écoles** des informations sur les infrastructures, l'emplacement et les opinions des directeurs d'école sur la disponibilité des ressources à l'école et sur la manière dont ils interagissent avec les parents. Il existe une certaine variation dans l'objectivité des données collectées.

Tableau 6. Questionnaires typiques et données collectées dans les évaluations transnationales

Questionnaire	Domaines de questions typiques
Élève	Informations démographiques (sexe, âge) Contexte familial et socio-économique Expériences liées à l'école (y compris l'exposition à l'intimidation) Expériences liées à l'apprentissage (activités en classe) Perceptions de soi, intérêts et aspirations liés à différents sujets Utilisation et maîtrise des technologies de l'information et de la communication (TIC)
Enseignant	Informations démographiques et générales (sexe, âge, années d'enseignement, matières enseignées) Diplômes et formation Types de pratiques pédagogiques utilisées et défis rencontrés
Directeur d'école	Caractéristiques de l'école Opinions sur la disponibilité et l'adéquation des ressources Gestion et gouvernance Interaction avec les parents et les communautés scolaires

Certains de ces questionnaires fournissent des informations sur certains indicateurs de l'ODD 4. Ces questionnaires déterminent également pour quelles sous-populations ces indicateurs peuvent être estimés sur la base de caractéristiques telles que la situation urbaine ou rurale des écoles, le statut socio-économique des élèves (par rapport aux autres élèves, et non à la

population) et le sexe des élèves et des enseignants, le tout qui peuvent éclairer les discussions sur l'équité.⁶

L'élaboration de conseils politiques présente des défis importants, en raison des différences dans les définitions des questionnaires selon diverses dimensions, telles que les classifications rurales/urbaines, le statut socio-économique (ou la richesse), la période de référence, etc. Les différences dans les définitions entravent la comparabilité des résultats de l'évaluation et l'interprétation des données. Par exemple, ce qui constitue « rural » dans un pays peut être différent dans un autre. De même, différentes définitions du statut socio-économique peuvent avoir un impact sur l'analyse des disparités dans les résultats scolaires selon les différentes couches sociales ou économiques. Comblar ces lacunes définitionnelles nécessite une collaboration internationale et le développement de cadres standardisés qui garantissent l'uniformité des définitions, permettant des évaluations précises et facilitant des comparaisons significatives entre les régions et les contextes socio-économiques.

FUTUR AGENDA

Les solutions suivantes sont proposées pour les problèmes identifiés dans ce document.

Manuel d'harmonisation des évaluations et de rapport

Avec les progrès réalisés ces dernières années, il est temps de compiler et de mettre régulièrement à jour un manuel contenant toutes les informations sur les critères d'éligibilité pour le rapport. Ce manuel servira de guide technique pour les programmes d'évaluation des apprentissages et présentera et mettra à jour le menu d'évaluations pouvant faire l'objet de rapports. Les aspects à inclure sont :

⁶Pour plus d'informations, consulter l'UIS (2022). *Monitoring of the Sustainable Development Goals using large-scale international assessments*
<https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2022/04/Monitoring-of-the-SDGs-Using-Large-Scale-International-Assessments-April-2022.pdf>.

- Alignement sur les normes et cadres d'apprentissage (les évaluations doivent mesurer de manière adéquate le programme d'études et les compétences prévus) ;
- Propriétés psychométriques (preuves de fiabilité, de validité, de difficulté appropriée et de discrimination, entre autres).
- Représentativité (les échantillons doivent refléter les populations cibles.)
- Comparabilité des administrations (des procédures d'administration cohérentes et standardisées affectent la comparabilité)
- Transparence des processus (la conception de l'évaluation, l'échantillonnage et l'analyse doivent être bien documentés)
- Capacité de liaison (suffisamment d'éléments/niveaux de compétence équivalents pour permettre la liaison)
- Implication des parties prenantes (processus consultatif tout au long de la conception et de la mise en œuvre)
- Faisabilité de la participation (coûts raisonnables, calendriers, renforcement des capacités et charges pour les pays)

Systeme d'accreditation

C'est donc également le bon moment pour introduire un système d'accréditation clair et transparent. Les prestataires d'évaluation, y compris les organisations gouvernementales, pourront demander à ce que les évaluations soient vérifiées quant à leur adéquation à l'objectif de rendre compte de l'indicateur 4.1.1 des ODD. Sur la base du manuel, une liste de contrôle contiendra les normes et les critères d'éligibilité auxquels les candidats doivent se conformer.

Augmenter les investissements et l'approche pour financer les données d'apprentissage

La couverture des données reste inégale, en particulier pour les pays en développement. Il est essentiel d'investir davantage pour étendre les données d'évaluation des apprentissages à l'échelle mondiale, dans le cadre d'une approche holistique visant à créer une infrastructure pour la production de données au niveau national, en transférant les connaissances et les compétences, mais également le pouvoir d'achat et la prise de décision.

Promouvoir la standardisation des questionnaires contextuels dans les évaluations des apprentissages

Les questionnaires contextuels constituent une prochaine étape importante pour soutenir la comparabilité. Un accord sur les définitions clés et un accord sur les éléments et le format standard pour capturer les caractéristiques individuelles des élèves, des enseignants et des écoles permettraient une meilleure comparabilité de l'équité et des facteurs déterminants des dimensions d'apprentissage.

Piloter l'approche mini-LAMP pour mesurer l'alphabétisation des adultes afin d'accroître la couverture de l'indicateur ODD 4.6.1

La faible couverture de l'indicateur mondial 4.6.1 des ODD signifie qu'il sera supprimé de la liste des indicateurs mondiaux lors de la révision de 2025 par le Groupe interinstitutions et d'experts sur les indicateurs des ODD. Pourtant, les compétences des adultes resteront à l'ordre du jour de l'éducation mondiale et une solution rentable demeure une priorité. Les États membres sont invités à tester l'outil mini-LAMP, qui peut être ajouté aux enquêtes existantes auprès des ménages.

Méthodologies innovantes sur les indicateurs à faible couverture

Les ODD 4.7.4 et 4.7.5 méritent l'exploration d'approches reflétant l'approche AMPL de 4.1.1 en permettant un module qui mesure les normes minimales convenues en tirant parti des évaluations existantes, telles que l'étude internationale sur l'éducation civique et la citoyenneté de l'IEA (ICCS).