

Guide rapide No. 3

Mettre en œuvre une évaluation nationale de l'apprentissage



Organisation
des Nations Unies
pour l'éducation,
la science et la culture



INSTITUT
DE STATISTIQUE
DE L'UNESCO



OBJECTIFS
DE DÉVELOPPEMENT
DURABLE



PARTENARIAT
MONDIAL pour
L'ÉDUCATION
*une éducation de qualité
pour tous les enfants*





Table des matières

Introduction.....	4
1. Quelles décisions doivent être prises avant de lancer l'évaluation ?.....	7
2. Comment mesurer l'apprentissage ?	11
3. Comment sélectionner les écoles et les élèves ?.....	15
4. Comment administrer l'évaluation ?.....	17
5. Comment créer une base de données ?.....	19
6. Comment calculer les résultats ?	21
7. Comment communiquer les résultats ?	22
8. Quels sont les besoins en termes de personnel et d'installations ?	24
9. Comment établir un budget ?.....	26
10. Quels sont les principaux problèmes et défis ?	28
Où trouver davantage d'informations ?.....	31

Publié en 2019 par :

Institut de statistique de l'UNESCO
P.O. Box 6128, Succursale Centre-Ville
Montréal, Québec H3C 3J7 Canada

Tél. : +1 514-343-6880
Courriel : uis.publications@unesco.org
<http://uis.unesco.org>

Préparé par : María José Ramírez, consultante, mariajose.ramirez001@gmail.com
ISBN 978-92-9189-239-6
Réf. : UIS/2017/LO/TD/16

© UNESCO-UIS 2019

Introduction



Apprendre est fondamental, d'une part pour les personnes car cela leur permet de tirer pleinement parti de leur potentiel, et d'autre part pour le développement social et économique des pays. Par conséquent, il est important pour les pays de savoir dans quelle mesure les élèves apprennent.

Pour le découvrir, de solides évaluations des apprentissages doivent être mises en place. Les pays mettent en œuvre des évaluations nationales afin de surveiller et de soutenir l'apprentissage pour tous. Pour ce faire, ils administrent des tests d'évaluation afin de mesurer ce que savent les élèves et ce qu'ils peuvent faire de leurs connaissances. Les résultats des évaluations sont ensuite utilisés pour alimenter les politiques et pratiques éducatives.

L'objectif de ce Guide rapide est d'aider les pays à mettre en œuvre leurs évaluations nationales de l'apprentissage. À cette fin, il répond aux questions suivantes :

- Quelles décisions doivent être prises avant de lancer l'évaluation ?
- Comment mesurer l'apprentissage ?
- Comment sélectionner les écoles et les élèves ?
- Comment administrer l'évaluation ?
- Comment créer une base de données ?
- Comment calculer les résultats ?
- Comment communiquer les résultats ?
- Quels sont les besoins en termes de personnel et d'installations ?
- Comment établir le budget ?
- Quels sont les principaux problèmes et défis ?

Élaborer une évaluation d'apprentissage nationale est un travail de longue haleine qui nécessite une volonté politique, un financement stable et des capacités techniques locales. Les pays qui mettent en œuvre une évaluation nationale de l'apprentissage pour la première fois peuvent apprendre beaucoup de choses des expériences d'autres pays. Le Mozambique est un pays à faible revenu qui a parfaitement réussi à mettre en œuvre graduellement son évaluation nationale de l'apprentissage (*voir Encadré 1*).

En mettant en place de solides évaluations de l'apprentissage, les pays seront davantage en mesure de soutenir l'apprentissage pour tous.

ENCADRÉ 1. L'ÉVALUATION NATIONALE DE L'APPRENTISSAGE AU MOZAMBIQUE

Le Mozambique est un pays à faible revenu du sud-est de l'Afrique. Après avoir considérablement augmenté le nombre des écoles à travers le pays, le Mozambique a focalisé son attention sur l'apprentissage. Ce faisant, il s'est intéressé à l'élaboration d'une évaluation dont l'objectif est de surveiller les tendances de l'apprentissage dans les domaines clés du programme.

La première évaluation nationale a été administrée en 2013. Un échantillon d'élèves de 3^{ème} année représentatifs au plan national (6 764 élèves de 400 écoles) a subi un test de lecture du portugais (la langue d'enseignement). Le test a mesuré la capacité des élèves à lire des lettres, des syllabes, des mots, de simples phrases et de courts paragraphes. L'administration du test était normalisée. Les administrateurs externes ont lu à haute voix les instructions du test ainsi que la plupart des questions devant des classes entières d'élèves. Les enseignants et les directeurs d'écoles ont répondu à des questionnaires contextuels.

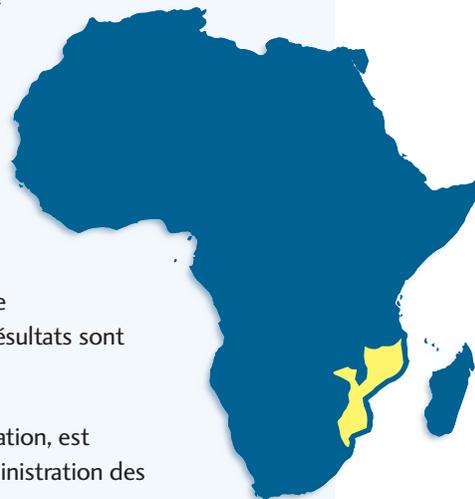
En 2016, l'évaluation a été administrée une nouvelle fois, en lecture et en mathématiques. Une nouvelle évaluation est prévue pour 2019.

Les résultats de l'évaluation montrent le pourcentage d'élèves par niveaux de compétence ; le niveau le plus élevé (Niveau III) correspond aux attentes du programme d'études. Ils montrent également les différences de résultats entre 2013 et 2016. Les résultats sont publiés dans un rapport national.

INDE, un institut de recherche semi-indépendant qui fait partie du ministère de l'Éducation, est responsable de l'évaluation nationale. INDE est en charge de la conception et de l'administration des tests, du traitement des données, de leur analyse et de l'établissement du rapport.

INDE a une petite équipe qui a graduellement renforcé ses capacités techniques relatives à l'évaluation. L'équipe disposait d'une expérience préalable, elle avait mis en œuvre l'évaluation internationale SACMEQ et réalisé le projet d'évaluation d'une classe de formation (Quizz I Can Read), ce qui lui permit d'acquérir des bases pour élaborer l'évaluation nationale. L'équipe a également bénéficié d'une participation à un programme de diplôme de master, en effectuant des visites d'études et une formation pratique (avec le soutien du READ Trust Fund).

Les frais de l'évaluation nationale ont été couverts par le gouvernement et par un fonds commun de donateurs.



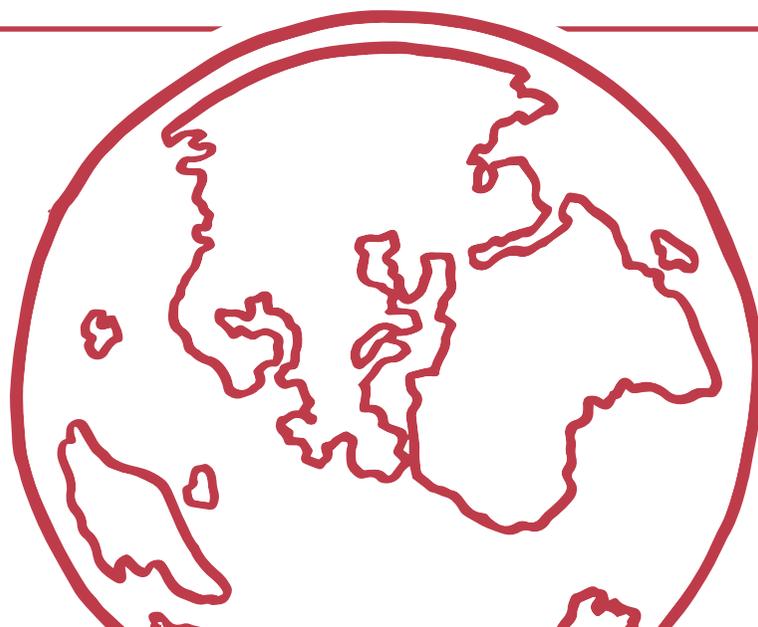
Quelles décisions doivent être prises avant de lancer l'évaluation ?

- **QUEL est l'objectif de l'évaluation ?** Les pays élaborent des évaluations de l'apprentissage pour surveiller dans quelle mesure les élèves atteignent les principaux objectifs d'apprentissage spécifiés dans le programme national et pour soutenir l'apprentissage pour tous. Certains pays utilisent également les résultats des évaluations pour tenir les écoles responsables et fournir aux élèves et à leurs parents des informations sur les progrès réalisés dans le cadre de l'apprentissage. Il est essentiel d'identifier les questions d'ordre politique auxquelles l'évaluation doit répondre (par exemple, Quel pourcentage d'élèves doit atteindre les objectifs du programme en mathématiques à la fin du primaire ? Les résultats s'améliorent-ils au fil du temps ?). L'évaluation doit être conçue de façon à permettre de répondre à ces questions.
- **QU'EST-CE qui sera testé ?** Les évaluations mesurent habituellement les domaines du programme (domaines d'apprentissage) considérés comme fondamentaux pour réussir dans la vie, tels que la compréhension en lecture, les mathématiques ou les sciences. D'autres mesurent également des domaines innovants tels que l'écriture, la santé ou les finances de manière à obtenir une image de l'apprentissage des élèves. Il est également important de décider dans quelle(s) langue(s) l'évaluation sera administrée.
- **QUELLES autres informations seront-elles collectées ?** Les pays doivent décider s'ils vont utiliser des questionnaires pour collecter les informations relatives aux variables contextuelles qui sont importantes pour comprendre l'apprentissage (par exemple, les méthodes d'enseignement et la disponibilité des manuels scolaires). Par exemple, plusieurs pays présentent des questionnaires aux élèves, à leurs enseignants, aux chefs d'établissement et aux parents.



- **QUI sera testé ?** Généralement, les pays sélectionnent une ou plusieurs années cibles (par exemple, la dernière année du primaire) pour l'évaluation. En fonction de la façon dont un pays souhaite utiliser les résultats de l'évaluation, il peut avoir besoin de tester un échantillon d'élèves représentatif au niveau national ou tous les écoles et élèves (recensement). Dans les deux cas, les pays ont besoin de disposer de la liste complète de toutes les écoles afin de pouvoir échantillonner de façon appropriée ou d'évaluer l'ensemble de leurs élèves. Dans les pays où le pourcentage d'enfants et d'adolescents est élevé, il peut être souhaitable de les tester aussi en utilisant les sondages auprès des ménages.
- **COMMENT les élèves seront-ils testés ?** Actuellement, la majorité des évaluations de l'apprentissage sont administrées sur papier et utilisent à la fois des questions à choix multiples et des questions ouvertes. Certains pays administrent des types de tests similaires, mais en ligne. La plupart des pays procèdent à des administrations groupées (par exemple, toute la classe), mais d'autres font passer les tests aux élèves individuellement. Les administrateurs des tests doivent suivre des directives précises afin de s'assurer que tous les élèves sont testés dans des conditions uniformes (c'est-à-dire dans le cadre d'une administration normalisée).
- **QUI administrera l'évaluation ?** Des administrateurs externes peuvent être formés pour administrer les tests à tous les élèves dans des conditions identiques. Des enseignants peuvent également s'en charger s'ils respectent les directives de l'administration.

Chaque pays répondra à ces questions en fonction de ses objectifs plus larges en matière d'éducation, du contexte local et des ressources disponibles.



- **QUAND l'évaluation sera-t-elle administrée ?** Certains pays administrent l'évaluation tous les ans, d'autres tous les deux ans, voire encore moins fréquemment. Certains pays varient l'année testée (par exemple, la 3^{ème} année est testée une année, la 6^{ème} année l'année suivante). La fréquence de l'évaluation variera en fonction de l'objectif de l'évaluation et du financement. Certains pays administrent l'évaluation au début de l'année scolaire, d'autres à la fin. Les dates de l'évaluation doivent tenir compte des principaux événements nationaux et de l'année scolaire. L'évaluation doit être administrée pendant la même période à chaque fois de façon à ce que les résultats soient comparables au fil du temps.
- **COMMENT les résultats seront-ils publiés ?** Idéalement, les résultats sont publiés dès que possible. Ils doivent répondre aux principales questions politiques (par exemple, Est-ce que les garçons et les filles ont des résultats similaires en lecture ?). Ils peuvent être publiés de façon à faire ressortir les pourcentages d'élèves atteignant les différents niveaux de connaissance (par exemple, Avancé, de Base, Insuffisant), les résultats moyens, le pourcentage de réponses correctes, ou d'autres éléments. Publier et analyser fréquemment les résultats permet d'informer différents publics et parties prenantes de la façon dont le pays tente de relever les défis pédagogiques. Cela permet également d'inspirer confiance dans l'évaluation nationale.

Une évaluation nationale des apprentissages peut coûter entre 200 000 et 1 000 000 de dollars des Etats-Unis, mais le coût varie en fonction de plusieurs facteurs.

- **COMBIEN cela va-t-il coûter ?** Une évaluation nationale des apprentissages peut coûter entre 200 000 et 1 000 000 de dollars des Etats-Unis, mais le coût varie en fonction de plusieurs facteurs, par exemple, la population cible (les enfants et les adolescents scolarisés par opposition aux enfants et adolescents non scolarisés), le nombre d'élèves testés, le mode d'administration (par exemple, une administration groupée ou individuelle), les coûts locaux des services (par exemple, l'impression) et le personnel (par exemple, les administrateurs des tests). Il est important d'estimer les coûts totaux et de prévoir un financement suffisant et stable (par exemple, obtenu auprès du gouvernement et de donateurs).
- **QUI va mettre en œuvre l'évaluation ?** Cela peut être le ministère ou le département de l'Éducation, un organisme gouvernemental semi-autonome ou une université disposant d'un groupe spécialisé dans les évaluations. Il est important de spécifier quels membres du personnel et quelles installations de l'unité seront disponibles pour l'évaluation. L'unité doit être responsable devant un organisme clairement identifiable (par exemple, un département du gouvernement).
- **QUI aura la charge de prendre ces décisions clés ?** Il est recommandé qu'un comité de pilotage prenne les décisions clés sur ce qui doit être évalué, par qui, auprès de qui et avec quelle fréquence. Le comité peut être placé sous l'égide du système éducatif national et composé de représentants des principales parties prenantes (par exemple, les responsables de l'élaboration des politiques, les pédagogues, les parents et les spécialistes des évaluations). L'implication de parties prenantes extérieures au système éducatif national peut contribuer à garantir un engagement national et une meilleure compréhension des questions relatives à l'éducation et aux évaluations. Des experts techniques peuvent conseiller le Comité de pilotage.

Chaque pays répondra à ces questions en fonction de ses objectifs plus larges en matière d'éducation, du contexte local et des ressources disponibles. Les décisions prises à la lumière des réponses à ces questions doivent rejaillir sur un cadre de l'évaluation qui doit être mis à la disposition du public. Le cadre de l'évaluation spécifie les principales caractéristiques de l'évaluation nationale de l'apprentissage, guide la conception et la mise en œuvre de l'évaluation, sert à informer les parties prenantes (par exemple, les responsables de l'élaboration des politiques, les pédagogues, les enseignants) et met en exergue l'approche de l'évaluation qui doit être adoptée par le pays.

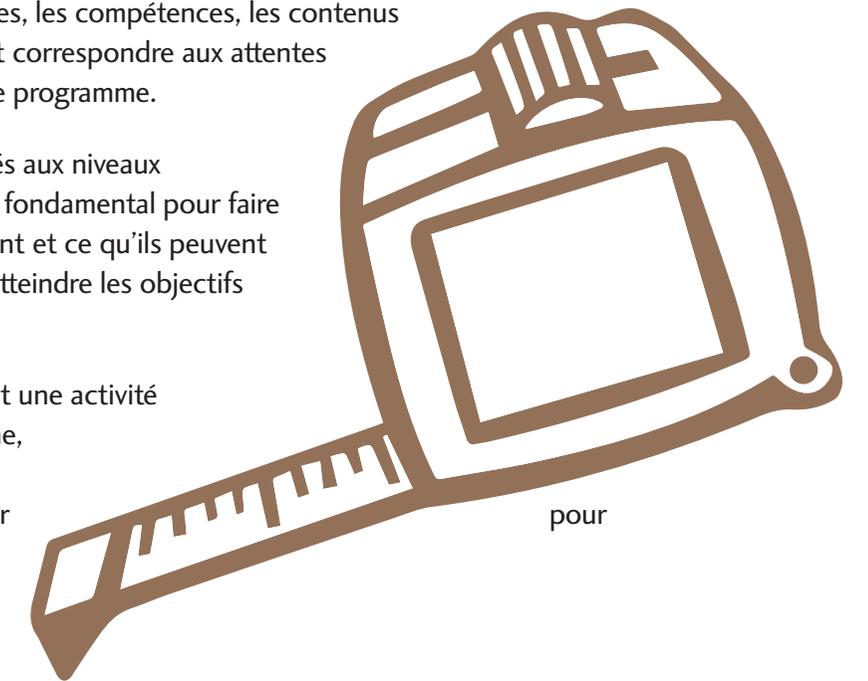
Comment mesurer l'apprentissage ?

Généralement, les pays mettent en œuvre les évaluations des apprentissages pour vérifier si les élèves atteignent les principaux objectifs pédagogiques et pour soutenir l'apprentissage pour tous. Les principaux objectifs pédagogiques sont généralement mentionnés dans le programme national.

Les tests des évaluations doivent être alignés avec le programme de façon à ce que leurs résultats puissent être interprétés et indiquent si les élèves ont atteint (ou n'ont pas atteint) les objectifs pédagogiques. En d'autres termes, les compétences, les contenus et les aptitudes mesurés par les tests doivent correspondre aux attentes en matière d'apprentissage spécifiées dans le programme.

Les tests des évaluations doivent être adaptés aux niveaux d'apprentissage effectifs des élèves. Ceci est fondamental pour faire rapport sur ce que les élèves savent réellement et ce qu'ils peuvent faire, et pour savoir s'ils sont loin ou près d'atteindre les objectifs pédagogiques.

Il est essentiel que l'élaboration des tests soit une activité collaborative entre des experts du programme, des enseignants et des spécialistes des évaluations. Les étapes importantes à franchir pour mesurer l'apprentissage sont les suivantes :



pour

- **Identifier les axes, compétences, objectifs, contenus ou aptitudes principaux.** Par exemple, la maîtrise de la lecture peut être structurée autour de deux axes : le décodage et la compréhension de la lecture. Les mathématiques peuvent être structurées autour des types de contenus suivants : nombres, géométrie, mesure et algèbre. Cette étape a des implications en termes de format des évaluations. Par exemple, un test effectué sur papier permet de mesurer la compréhension de la lecture, mais pas la compréhension orale. Dans la mesure du possible, le format de l'évaluation doit ressembler au style des connaissances et compétences acquises par les élèves dans leurs classes.
- **Identifier les principaux niveaux de connaissance.** Chaque niveau (par exemple, Avancé, de Base, Insuffisant) décrit ce que savent et peuvent faire les élèves aux différents stades de la trajectoire de leur apprentissage. Par exemple, le plus haut niveau de connaissance peut décrire les objectifs pédagogiques, alors que le niveau de connaissance le plus faible peut décrire ce que les élèves ayant obtenu les moins bons résultats peuvent réellement faire. Les résultats de l'évaluation peuvent être présentés dans un rapport montrant le pourcentage d'élèves ayant atteint chacun des niveaux.
- **Écrire les spécifications du test pour chacun des domaines d'apprentissage.** Identifier les principales catégories d'analyse afin de guider l'élaboration du test : axes, compétences, objectifs, contenus ou aptitudes : formats des questions (par exemple, questions à choix multiples ou questions ouvertes) ; type de texte (par exemple, narratif, informatif) ; et complexité ou difficulté des questions (par exemple, facile, moyenne, difficile). Spécifier le nombre de rubriques ou de questions qui seront nécessaires, globalement et dans chaque catégorie. La répartition des questions doit être cohérente avec le programme.



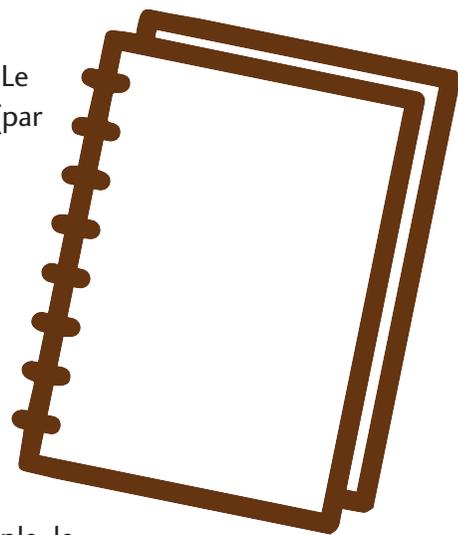
- **Recruter et former les auteurs des sujets.** Les auteurs des sujets doivent avoir une bonne compréhension du cadre de l'évaluation et des spécifications du test avant de commencer à écrire les sujets.

Les auteurs des sujets peuvent faire partie de l'organisation responsable de l'évaluation nationale ou être extérieurs à cette unité (par exemple, des enseignants ou des spécialistes du programme appartenant à une autre unité du ministère ou du département de l'Éducation).

- **Écrire les éléments du test.** Les éléments (par exemple, les questions ou les tâches) doivent couvrir un large éventail de niveaux de difficulté afin de mesurer les différents niveaux d'apprentissage. Les bons éléments (a) présentent des tâches qui peuvent être rapprochées des spécifications du test ; (b) sont conceptuellement corrects ; (c) sont focalisés sur des questions importantes, et non pas sur des détails triviaux ; et (d) expliquent aux élèves ce qu'ils doivent faire. Les bons éléments à choix multiples (a) comprennent également une clé qui est indéniablement correcte ; (b) comprennent des fausses pistes qui sont indéniablement incorrectes, tout en étant raisonnables et plausibles ; et (c) présentent des options de réponses qui sont équivalentes en termes de longueur et de style. Les questions ouvertes doivent être écrites en même temps que les guides relatifs à leurs résultats et doivent être notées après l'administration du test.
- **Écrire les instructions du test.** Les instructions doivent accueillir les élèves, leur indiquer comment effectuer le test et en combien de temps et fournir quelques exemples pratiques. Les élèves doivent comprendre comment répondre au test, y compris aux questions à choix multiples. Si une feuille distincte est fournie pour les réponses, les élèves doivent comprendre comment indiquer leurs réponses sur la feuille. Les instructions doivent être lues par l'administrateur du test avant que les élèves ne commencent à répondre au test.
- **Réviser et pré-piloter les éléments.** Les éléments doivent être révisés par les auteurs des éléments ainsi que par des groupes de personnes extérieures qui n'ont pas été impliquées dans la rédaction des éléments. Les éléments révisés doivent être pré-pilotés en les administrant à un petit nombre d'élèves (par exemple, 50 élèves). Les examens et révisions ultérieurs doivent être effectués avant d'assembler les brochures du test pilote.
- **Procéder à une administration pilote.** Ceci est important pour identifier les caractéristiques psychométriques (par exemple, la difficulté ou la discrimination) de l'ensemble du test et de chaque élément. Pour le test pilote, une bonne pratique consiste à inclure deux ou trois fois le nombre d'éléments requis pour le test final et à administrer chaque élément à au moins 200 élèves.

ayant les mêmes caractéristiques que ceux qui subiront le test final. Le pilote est également important pour tester le manuel, la logistique (par exemple, les transports) et d'autres procédures de l'administration.

- **Analyser les résultats du pilote.** Calculer les statistiques des éléments tels que la difficulté (par exemple, le pourcentage de réponses correctes, les scores gradués) et la discrimination (par exemple, les corrélations point-bisériales) ; le pourcentage d'élèves par option de réponse (pour les questions à choix multiples) ou le pourcentage d'élèves ayant obtenu une reconnaissance totale ou partielle (pour les questions ouvertes) ; et le pourcentage omis. Les principales statistiques des tests concernent les scores moyens, la fiabilité et la précision (par exemple, le coefficient alpha de Cronbach). Inclure des éléments couvrant un large éventail de difficultés et obtenir un coefficient alpha de Cronbach supérieur à 0,70 est considéré comme une bonne pratique. Le logiciel IATA est gratuit et peut être utilisé pour procéder à ces analyses.
- **Assembler les brochures du test en vue de la principale administration du test.** Les brochures du test doivent être plaisantes sans être distrayantes. Elles doivent comprendre des numéros d'identification uniques permettant de localiser les élèves et les écoles. Le nombre d'éléments présents dans les brochures doit être approprié de façon à ce que les élèves puissent répondre au test dans un délai raisonnable. Il est bon de placer les éléments les plus faciles au début des brochures. Des pays utilisent souvent différentes brochures comprenant différents éléments, afin d'augmenter le nombre global d'éléments testés. Dans ce cas, les différentes brochures doivent être équivalentes en termes de contenus et de caractéristiques psychométriques (par exemple, la difficulté). Elles peuvent également partager certains éléments communs.
- **Préparer l'impression.** L'impression est onéreuse et peut prendre plusieurs semaines. Des devis doivent être demandés dès que les principaux paramètres sont connus (par exemple, le nombre de brochures à imprimer et le nombre de pages par brochure). Tous les brochures des tests et autres instruments (par exemple, les questionnaires et manuels d'administration) doivent être scrupuleusement vérifiés avant l'impression.



Comment sélectionner les écoles et les élèves ?

Les écoles et les élèves doivent être sélectionnés de façon à ce que les résultats globaux de l'évaluation puissent être interprétés comme appartenant au pays considéré dans son ensemble. Dans certains cas, la sélection des écoles et des élèves doit être réduite en raison de contraintes financières ou logistiques. Dans ces cas, les résultats de l'évaluation doivent être interprétés comme appartenant à un niveau infranational du pays (par exemple, le pourcentage d'élèves atteignant le niveau de connaissance minimum dans les écoles urbaines d'une région X). Toutes les décisions relatives à la sélection des écoles et des élèves doivent être justifiées.

Les principales étapes de la sélection des écoles et des élèves en vue de l'évaluation de l'apprentissage sont les suivantes :

- **Spécifier la population cible des élèves.** Indiquer à quels élèves et écoles du pays se réfèrent les résultats (par exemple, tous les élèves en 6^{ème} année, des écoles publiques et privées).
- **Construire ou obtenir la liste complète des écoles.** Cette liste (base d'échantillonnage) devrait idéalement inclure des informations d'identification et de contact pour toutes les écoles cibles, des informations relatives à la classification des écoles (par exemple, rurales/urbaines, publiques/privées, région/district), le nombre de



classes et le nombre d'élèves dans l'année cible.

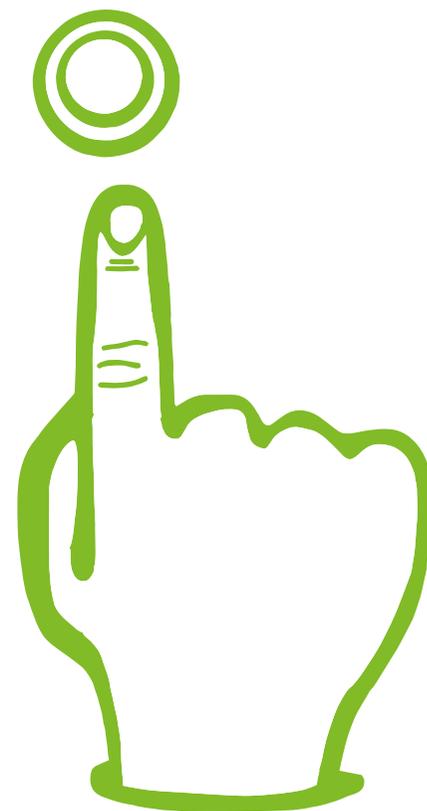
- **Décider de la meilleure approche pour sélectionner les écoles et les élèves :**
 - a. Sélectionner un échantillon d'élèves qui représente le niveau national (ou infranational). C'est la meilleure approche à adopter si l'objectif de l'évaluation est d'estimer les niveaux d'apprentissage dans l'ensemble du pays. Les méthodes d'échantillonnage impliquent généralement i) une stratification de l'échantillonnage en fonction des informations relatives à la classification des écoles (par exemple, par région) ; un échantillonnage des écoles aléatoire et systématique, à partir de la base d'échantillonnage, avec les grands établissements ayant davantage de chances d'être sélectionnés ; et iii) l'échantillonnage d'une classe ou d'élèves intacts, au sein des écoles. C'est également la meilleure pratique à adopter pour sélectionner des écoles de remplacement dans les cas où certaines des écoles sélectionnées ne peuvent pas participer.
 - b. Sélectionner tous les écoles et élèves (par exemple, recensement) dans l'année cible. C'est généralement la meilleure approche à adopter pour établir des rapports sur les résultats au niveau des écoles ou si l'évaluation est mise en œuvre dans un pays relativement petit.



Comment administrer l'évaluation ?

L'administration de l'évaluation de l'apprentissage est une initiative majeure. Elle exige généralement de contacter des centaines d'écoles, de recruter et de former des centaines d'administrateurs et de superviseurs de test, de distribuer des centaines de brochures de test partout dans le pays en toute sécurité, et de sous-traiter des services (par exemple, de transport). Les opérations sur le terrain sont toutes les activités nécessaires pour que l'administration du test soit menée à bien. Elles consistent à :

- **Concevoir le plan des opérations sur le terrain.** Ce plan doit identifier les activités, les responsabilités et les calendriers. Il indique de quelle façon les écoles seront attribuées aux coordonnateurs sur le terrain et aux administrateurs de test (par exemple, cinq écoles par coordonnateur). Un manuel des opérations sur le terrain et une documentation les concernant (par exemple, des formulaires et des listes de contrôle) doivent présenter les lignes directrices et être utilisés pour contrôler la qualité de l'administration.
- **Former les principaux membres du personnel.** La formation doit mettre l'accent sur les objectifs de l'évaluation, l'importance de conditions de test uniformes, ainsi que les rôles et les responsabilités

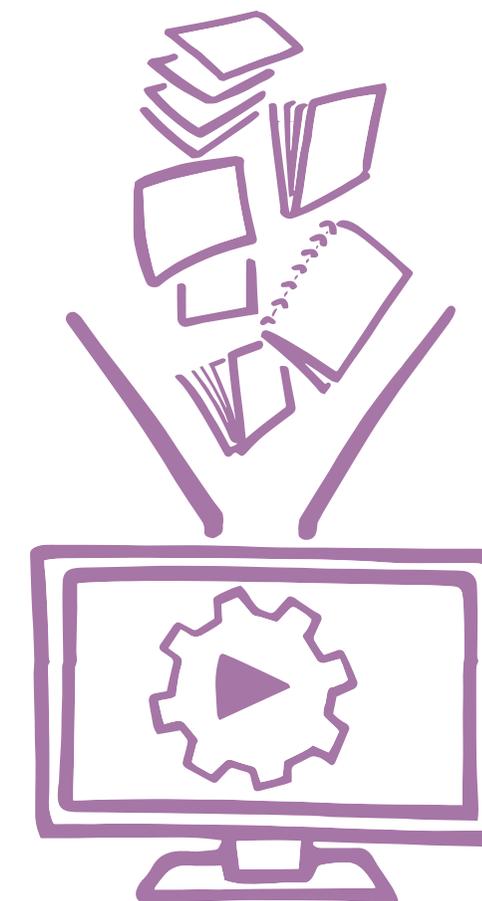


des formateurs pendant les opérations sur le terrain. Elle doit être dispensée avant la collecte des données et reproduire, dans la mesure du possible, les conditions réelles de l'administration du test, par exemple, en lisant à haut voix le manuel d'administration du test pendant les sessions de formation et en remplissant des formulaires concernant l'administration.

- **Contactez les écoles.** Inviter les écoles sélectionnées à participer à l'évaluation. S'assurer qu'il est possible de contacter une personne, à l'école, qui peut contribuer à l'organisation de l'administration du test (par exemple, en informant les enseignants de l'évaluation, en stockant les instruments du test, etc.).
- **Imprimer, vérifier et distribuer les instruments de l'évaluation.** Vérifier les échantillons des instruments imprimés (par exemple, les tests et les questionnaires) afin de s'assurer qu'ils satisfont aux normes de qualité. L'emballage et la distribution peuvent prendre plusieurs semaines. La sécurité et la confidentialité des instruments doivent faire partie des principales priorités.
- **Administrer l'évaluation conformément aux procédures normalisées.** Tous les élèves doivent subir le test dans les mêmes conditions, qui doivent être uniformes. Le manuel de l'administrateur du test doit stipuler ces conditions et servir de base à la qualité du contrôle.
- **Envoyer dans les écoles des observateurs du contrôle de la qualité.** Ceci est important pour vérifier que l'évaluation a été efficacement administrée dans toutes les écoles, dans les mêmes conditions uniformes. Il est également important de faire en sorte que l'évaluation inspire confiance. Les observateurs peuvent se référer au manuel d'administration ou à une liste de contrôle. Les observateurs peuvent comprendre des membres du personnel du ministère ou du département de l'Éducation, des donateurs, des parents ou d'autres personnes.
- **Vérifier le retour des instruments de l'évaluation.** Généralement, les pays ont besoin de collecter dans les écoles tous les brochures de test et les autres instruments (par exemple, les questionnaires et les formulaires). Tous les brochures et les instruments doivent être enregistrés (par exemple, en utilisant des numéros d'identification) et comptabilisés avant et après le test.

Comment créer une base de données ?

- **Préparer la structure d'une base de données afin d'enregistrer les données.** La base de données doit inclure toutes les variables des brochures de test et des autres instruments de l'évaluation (par exemple, les numéros d'identification, les réponses à l'élément 1, à l'élément 2, etc.). Elle doit spécifier le type d'informations à enregistrer pour chaque variable (par exemple, texte/numérique et plages valables). Ces informations doivent être documentées dans le manuel relatif à la base de données (c'est-à-dire le manuel de codage).
- **Préparer les protocoles du traitement des données.** Les protocoles de la saisie de données, du nettoyage des données et des vérifications. Par exemple, ils donnent des instructions aux responsables de l'entrée des données sur la façon dont elles doivent être saisies, énoncent les règles permettant de procéder aux contrôles de la cohérence interne (par exemple, si un élève est absent, toutes les réponses aux éléments doivent manquer) et des corrections automatiques.
- **Entrer les données brutes dans la base de données.** Les méthodes comprennent la collecte de données en ligne, la lecture optique des brochures de test ou des feuilles de réponses, et la

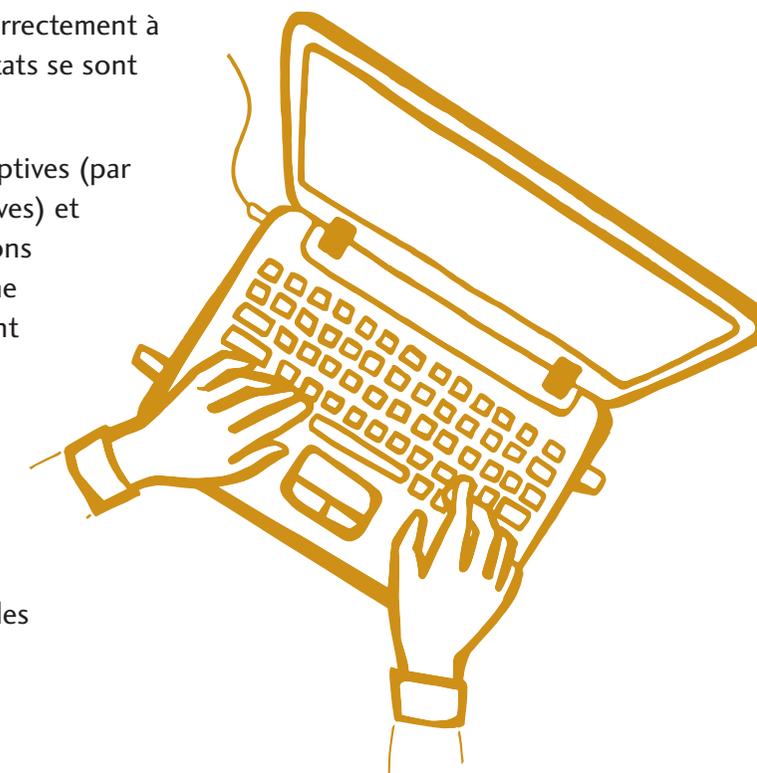


saisie manuelle. Des matériels (par exemple, des scanners ultra-rapides) et logiciels (par exemple, Access ou Excel) peuvent s'avérer nécessaires, selon la méthode utilisée. Des dispositions doivent être prises pour s'assurer de disposer des matériels et logiciels nécessaires pour accomplir cette tâche.

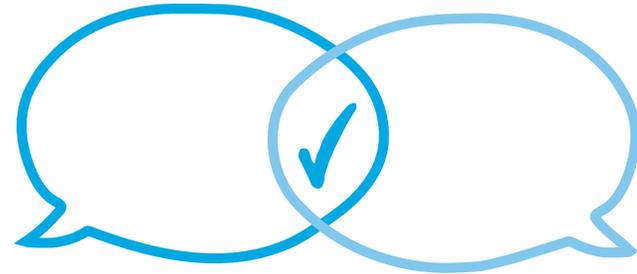
- **Nettoyer les données.** Cette étape comprend les procédures de vérification et de correction des numéros d'identification uniques, les dossiers en deux exemplaires ou abandonnés, et les vérifications visant à s'assurer que les valeurs saisies sont valables et conformes aux plages prévues. Le nettoyage des données doit généralement être effectué plusieurs fois avant de produire la base de données finale et officielle.
- **Calculer ou ajouter de nouvelles variables.** Elles peuvent inclure :
 - a. Scores gradués de la théorie de la réponse d'item. Ceux-ci sont utiles pour placer les scores de différentes brochures dans une même échelle et les comparer.
 - b. Niveaux de connaissance. Les élèves sont classés dans un niveau (Avancé, de Base, Insuffisant) sur la base de leur score final au test.
 - c. Pondérations visant à ajuster l'échantillonnage. Celles-ci sont importantes pour produire des résultats valables au niveau national.

Comment calculer les résultats ?

- **Identifier les principales questions politiques qui vont guider les analyses.**
Par exemple : Quel est le résultat moyen du pays en mathématiques ? Quel pourcentage d'élèves a atteint le niveau minimum de compétence en lecture ? Est-ce qu'une différence est perceptible entre les résultats des filles et des garçons ? Quel pourcentage d'élèves a répondu correctement à des éléments spécifiques du test ? Est-ce que les résultats se sont améliorés par rapport à l'évaluation précédente ?
- **Analyser les données.** Calculer les statistiques descriptives (par exemple les résultats moyens et les pourcentages d'élèves) et les statistiques inductives (par exemple, les comparaisons moyennes) afin de répondre aux questions de recherche les plus basiques. Reconnaissez que les résultats ne sont pas parfaitement précis (fiables) en faisant état des erreurs standards ou en indiquant si les différences sont statistiquement significatives.
- **Appliquer des pondérations** en tenant compte du plan d'échantillonnage.
- **Documenter toutes vos analyses** de façon à ce que les autres puissent les comprendre et les reproduire.



Comment communiquer les résultats ?



Les pays mettent en place des évaluations des apprentissages pour surveiller et soutenir l'apprentissage pour tous. Pour que cet objectif soit atteint, les résultats de l'évaluation doivent être communiqués efficacement aux parties prenantes.

Principales étapes à franchir pour assurer une communication efficace des résultats de l'évaluation :

- **Préparer un plan de communication.** Identifier les principales questions et messages politiques à communiquer, les produits de communication (par exemple, rapport national ou dépliant), les parties prenantes auxquelles chacun des produits sera destiné (par exemple, les enseignants, les parents), et le calendrier de la publication de ces produits (par exemple, trois mois après l'administration de l'évaluation). S'assurer de mobiliser suffisamment de fonds pour mettre en œuvre ce plan.
- **Identifier les principales questions et messages politiques.** Par exemple, combien d'élèves atteignent les objectifs du programme ? Y-a-t-il des différences entre les résultats régionaux et les résultats nationaux ? Est-ce que les résultats s'améliorent au fil du temps ?
- **Identifier les principaux publics ou parties prenantes :** Parents, enseignants, le ministère ou le département de l'Éducation, les organismes de financement, les médias et le grand public. S'assurer que leurs besoins en matière d'information sont satisfaits. Construire une culture de l'évaluation grâce à laquelle les parties prenantes comprennent, jugent et utilisent les informations de l'évaluation afin d'améliorer l'évaluation.
- **Sélectionner les meilleurs produits de communication.** Les différents produits (par exemple, les rapports, dépliants ou ateliers) servent différents objectifs et sont utiles pour sensibiliser les différents publics ou parties prenantes. S'assurer que les produits de communication sélectionnés répondent aux principales questions politiques (par exemple, Quel pourcentage d'élèves a atteint le

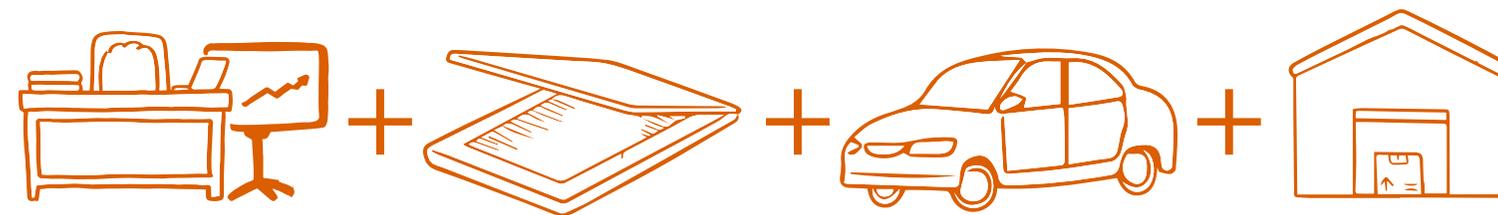
niveau minimum de compétence en lecture ?) et parviennent aux différentes parties prenantes (par exemple, en distribuant les rapports sur les écoles à l'ensemble des écoles). S'assurer que les informations mentionnées dans les rapports sont correctes et claires. Imprimer certains rapports peut nécessiter des bases de données supplémentaires (par exemple, une base de données des rapports sur les écoles qui permet d'imprimer les résultats spécifiques à chacune des écoles). Voici quelques exemples de produits de communication à prendre en considération :

- Les rapports nationaux destinés au grand public. Ils offrent une vue d'ensemble de l'évaluation et de ses résultats.
 - Les ateliers pour les enseignants et les chefs d'établissement. Donnent l'occasion d'expliquer l'importance de l'évaluation et de discuter de la façon d'utiliser les résultats pour améliorer l'apprentissage des élèves.
 - Les rapports à l'intention des médias. Contiennent toutes les informations nécessaires pour communiquer les résultats à un public plus large, y compris les principaux messages, graphiques et tableaux.
 - Les sites Web, vidéos, blogues et réseaux sociaux pour dialoguer régulièrement avec le grand public.
 - Les rapports à l'intention des écoles et les rapports à l'intention des parents sont des outils efficaces, notamment aux fins de responsabilisation.
 - Les conférences et les exposés (par exemple en format Power Point) donnent l'occasion de dialoguer avec les universités et les autres institutions (par exemple, les organismes de financement).
 - Les dépliants, brochures et posters peuvent être utilisés pour communiquer de brefs messages à un public plus large et à ceux qui n'ont pas le temps de lire les résultats plus élaborés (par exemple, les responsables de l'élaboration des politiques).
- **Présenter des mesures visant à améliorer les résultats de l'évaluation.** Ceci est particulièrement important lorsque l'on fait état de résultats insuffisants. Par exemple, le ministère de l'Éducation peut déclarer que ces résultats sont un diagnostic ou un point de départ, et lancer un nouveau programme pour soutenir l'apprentissage. Les écoles peuvent fixer des objectifs d'apprentissage à atteindre la prochaine fois que des résultats d'évaluation seront publiés.
 - **Surveiller la façon dont les résultats de l'évaluation sont utilisés.** Vérifier si les rapports parviennent comme prévu aux écoles, parents et aux autres parties prenantes. Rencontrer les parties prenantes et leur poser des questions sur les différents produits d'évaluation : Les ont-ils reçus ? Ont-ils trouvé que les résultats étaient intéressants pour leur école ? Ont-ils pu les comprendre ? Comment ont-ils utilisé les résultats pour soutenir l'apprentissage ?

Quels sont les besoins en termes de personnel et d'installations ?

Les principaux membres du personnel nécessaires pour mettre en œuvre l'évaluation nationale de l'apprentissage sont :

- **Coordonnateur principal.** Doit donner les instructions générales et diriger les opérations, s'assurer que l'évaluation répond aux principales questions politiques, qu'elle est techniquement appropriée, et qu'elle est effectuée en temps opportun et sans dépasser les limites du budget. Selon le pays, des coordonnateurs régionaux peuvent également s'avérer nécessaires. Le personnel administratif et le personnel spécialiste des technologies de l'information peuvent travailler directement auprès du coordonnateur national.
- **Coordonnateurs de l'élaboration des tests.** Ces personnes ont la responsabilité de produire les tests qui mesurent l'apprentissage dans chaque domaine d'apprentissage (par exemple, les mathématiques, la lecture). Ils doivent avoir une connaissance approfondie des objectifs du programme et du niveau d'apprentissage réel des élèves dans les classes. Les principales responsabilités comprennent le recrutement, la formation et le soutien des équipes d'auteurs des éléments et la notation des questions ouvertes.
- **Coordonnateur des opérations sur le terrain.** Est responsable de l'administration de tous les instruments d'évaluation auprès des élèves. Cette personne a la responsabilité de recruter et de former les administrateurs de test, les superviseurs, les coordonnateurs régionaux ou de district et les observateurs du contrôle de la qualité. Cette personne peut être employée à plein temps par l'organisation en charge de l'évaluation nationale, faire partie du personnel d'autres services du ministère ou du département de l'Éducation, ou peut être employé de manière temporaire. Le coordonnateur des



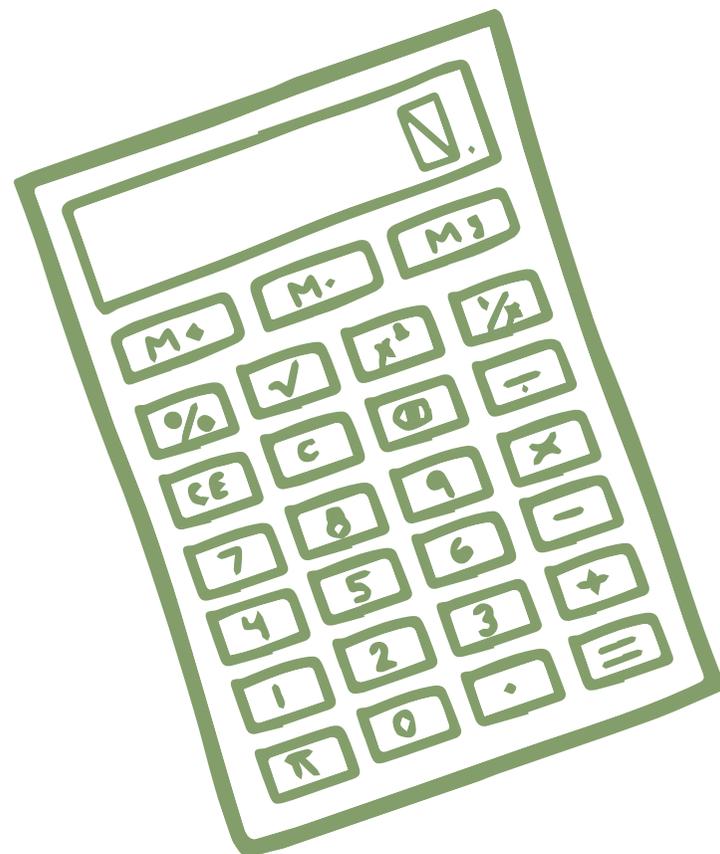
opérations sur le terrain est également responsable de la production du manuel d'administration et de la logistique de l'évaluation (par exemple, la distribution et la collecte des instruments).

- **Coordonnateur de la gestion des données.** Est responsable de la préparation des bases de données, c'est-à-dire de la saisie et du nettoyage des données, du calcul des pondérations, des résultats, des scores minimaux associés aux niveaux de connaissance et aux résultats. Cette personne doit s'assurer que les bases de données satisfont aux exigences relatives aux produits de communication. Cette personne doit avoir les compétences nécessaires pour exécuter les analyses statistiques et psychométriques à l'aide des différents logiciels (par exemple, SAS, SPSS, STATA et IATA).
- **Coordonnateur de la communication.** Est responsable de la mise en place et de la mise en œuvre d'un plan de communication efficace. Travailler avec un concepteur et un spécialiste de la communication est vivement recommandé.

L'organisation responsable de l'évaluation nationale doit pouvoir accéder aux installations de base, notamment aux :

- **Salles de réunion et bureaux** des principaux membres du personnel, avec des ordinateurs connectés à Internet, des imprimantes et les autres fournitures de bureau habituelles.
- **Matériels** (par exemple, scanners) **et logiciels** (par exemple, SPSS, IATA) avec les licences permettant de les utiliser.
- **Véhicules de transport** des cartons contenant les instruments de l'évaluation.
- **Installations de stockage sécurisées** qui contiennent de grandes quantités de cartons contenant les instruments de l'évaluation.

Comment établir un budget ?



Établissez un budget en identifiant les coûts des éléments principaux de l'évaluation et en estimant les coûts de chacun d'entre eux. Le Tableau 1 présente les coûts des principaux éléments ainsi que la ventilation proportionnelle des coûts, par élément. Les coûts totaux et la ventilation peuvent varier considérablement en fonction des pays, notamment en raison des caractéristiques de l'évaluation (par exemple, sondage ou échantillon), des capacités institutionnelles existantes et des coûts locaux. Vous remarquerez que le tableau ne comprend pas les coûts associés au personnel, aux bureaux et aux équipements.

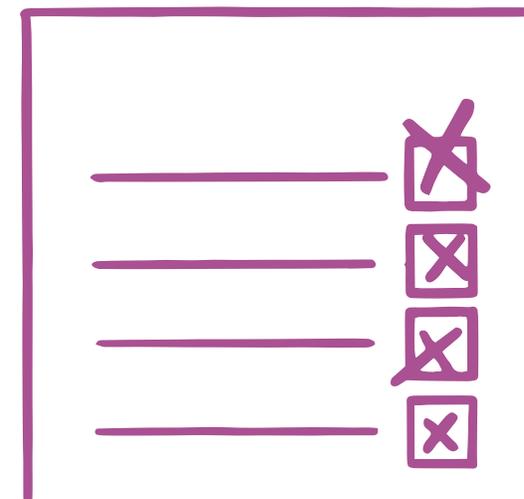
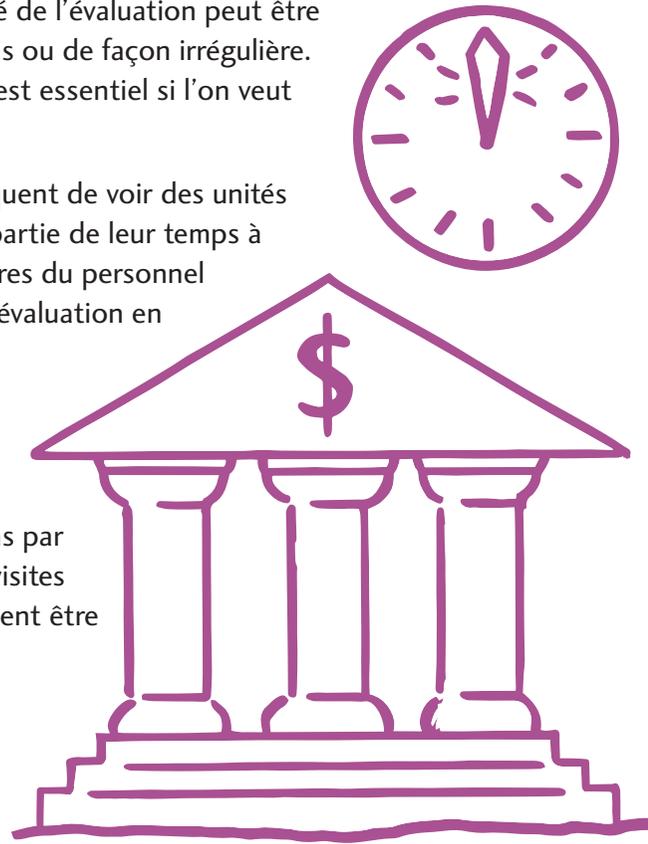
En cliquant ici, vous pouvez regarder comment les pays établissent le budget de leurs évaluations des apprentissages.

TABEAU 1. Coûts des principaux éléments et ventilation des coûts par élément d'une évaluation de l'apprentissage

ÉLÉMENT	VENTILATION DES COÛTS
CONCEPTION ET PILOTE DE L'ÉVALUATION	
Cadre de l'évaluation, spécifications du test	20%
Rédaction des sujets	
Élaboration du questionnaire	
Échantillonnage	
Impression du pilote	
Administration du pilote	
ADMINISTRATION PRINCIPALE	
Échantillonnage	50%
Brochures du test et conception des questionnaires	
Impression des instruments (brochures, questionnaires, manuels)	
Distribution dans les écoles	
Administration principale	
TRAITEMENT ET ANALYSE DES DONNÉES	
Saisie et nettoyage des données	15%
Notation des questions ouvertes	
COMMUNICATION	
Conception du site Web, des rapports, des dépliants et des posters	15%
Ateliers, conférences	
Distribution aux écoles, parents et autres parties prenantes	
TOTAL	100%

Quels sont les principaux problèmes et défis ?

- **Incapacité à mobiliser le soutien politique et un financement stable.** Les coûts de l'évaluation ne sont peut-être pas entièrement inclus dans le budget annuel du ministère ou du département de l'Éducation. Par conséquent, la stabilité de l'évaluation peut être menacée. L'évaluation peut n'être administrée qu'une seule fois ou de façon irrégulière. Impliquer les parties prenantes et rechercher la transparence est essentiel si l'on veut minimiser ce risque.
- **Nécessité de disposer d'un personnel suffisant.** Il est fréquent de voir des unités d'évaluation dont les membres ne peuvent consacrer qu'une partie de leur temps à l'évaluation. Il est important de convenir du nombre de membres du personnel et du temps qu'ils consacreront à l'évaluation et de planifier l'évaluation en fonction de ces paramètres.
- **Nécessité de renforcer les capacités locales.** La meilleure façon d'atteindre cet objectif consiste à dispenser une formation pratique tout en mettant en œuvre l'évaluation. Parmi les autres activités de renforcement des capacités, citons par exemple les cours ou ateliers (en face à face ou en ligne), les visites d'étude et la participation à des conférences. Des efforts doivent être déployés pour conserver le personnel formé.
- **Les tests sont trop difficiles.** Les tests d'évaluation sont généralement aptes à mesurer les objectifs ou les attentes du programme mais sont plus décevants lorsqu'il s'agit de



mesurer ce que peuvent faire les élèves ayant obtenu les résultats les plus faibles. Par conséquent, il peut arriver que les résultats finissent par montrer que la plupart des élèves n'arrivent pas à se hisser au-dessus des niveaux de connaissance les plus faibles. Il convient donc de s'assurer que la conception du test s'enrichit à la fois du programme et des niveaux de connaissance effectifs des élèves.

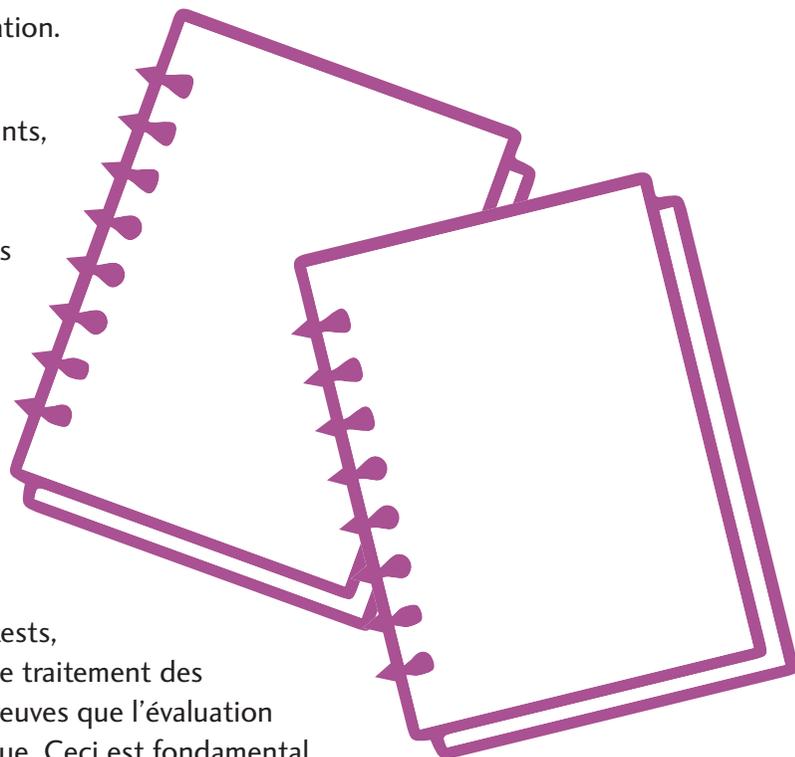
- **Les faiblesses de l'échantillonnage.** Les évaluations censées être représentatives au niveau national finissent parfois par utiliser des échantillons qui ne sont représentatifs que de quelques régions ou que d'un certain type d'écoles (par exemple, les écoles urbaines). Ou, pendant le travail sur le terrain, il peut arriver que seules les « écoles faciles à atteindre » subissent les tests. Ceci peut se produire en raison d'une mauvaise organisation, de difficultés d'ordre logistique ou d'un budget insuffisant. L'échantillonnage et le travail sur le terrain doivent être planifiés en détail et suffisamment à l'avance pour éviter ces problèmes.
- **L'absence de procédures normalisées.** Lorsqu'il n'y a pas de directives claires et de manuels pour administrer l'évaluation, lorsque les administrateurs des tests n'ont pas été formés de façon appropriée, lorsque les règles de la notation ne sont pas claires, lorsqu'il n'y a pas de règles ou de protocoles clairs pour créer l'ensemble de données, l'ensemble de l'effort d'évaluation peut être gâché. Un niveau de normalisation minimum est nécessaire pour s'assurer que les résultats des tests peuvent être interprétés conformément aux objectifs. Les procédures relatives aux manuels, à la formation et au contrôle de la qualité sont des outils importants pour vérifier la normalisation.
- **Les résultats d'évaluation ne sont pas comparables.** Ceci est un problème majeur lorsque l'objectif est de faire rapport sur les modifications apportées à l'apprentissage au fil des années. La comparabilité est menacée lorsque (a) les écoles et

les élèves échantillonnés ne sont pas équivalents de l'administration d'un test à l'autre ; et lorsque (b) les tests ne sont pas parallèles ou n'ont pas été placés sur la même échelle de notes. Un autre problème connexe est que l'incertitude (erreur) associée aux notes des tests n'est pas connue. Il est alors impossible de savoir si les différences apparaissant dans les notes sont importantes ou statistiquement significatives. Lors de la conception de l'évaluation, assurez-vous qu'elle a les caractéristiques techniques nécessaires pour permettre de comparer les résultats.

- **Les résultats de l'évaluation ne sont pas publiés.** Ceci peut très bien se produire dans les situations où les résultats sont médiocres. Il est important d'anticiper ce scénario avec le ministère de l'Éducation ou l'organisme équivalent. Par exemple, planifier une stratégie de communication lorsque les résultats médiocres sont utilisés comme point de départ de l'amélioration.

- **L'absence de culture de l'évaluation.** Les parties prenantes (par exemple, les enseignants, les parents et les responsables de l'élaboration des politiques) peuvent ne pas comprendre, apprécier et utiliser les résultats de l'évaluation comme prévu. Produisez des rapports, dépliants et sites Web simples qui abordent quelques questions de recherche essentiels et proposez des ateliers pour expliquer les résultats et l'évaluation en général.

- **L'absence de documentation.** La documentation technique (par exemple, le cadre de l'évaluation, les spécifications des tests, les plans d'échantillonnage, les protocoles de traitement des données) est nécessaire pour donner des preuves que l'évaluation peut être interprétée et utilisée comme prévue. Ceci est fondamental pour garantir la crédibilité de l'évaluation.



Où trouver davantage d'informations ?

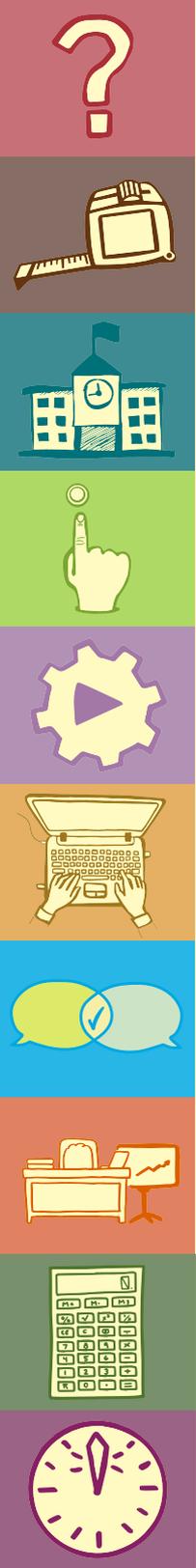
Ce Guide rapide s'appuie sur les références suivantes, qui couvrent les sujets présentés ici de façon beaucoup plus approfondie :

Banque mondiale (2008). *Évaluations nationales des acquis scolaires* (disponibles en anglais, espagnol, français, portugais et russe). Washington : Banque mondiale.

- Évaluer les niveaux nationaux de performance dans l'éducation (Vol. 1)
- Mettre au point les tests et questionnaires pour une évaluation nationale des acquis scolaires (Vol. 2)
- Mettre en œuvre une évaluation nationale des acquis scolaires (Vol. 3)
- Analyser les données issues d'une évaluation nationale des acquis scolaires (Vol. 4).
- Communiquer et utiliser les résultats d'une évaluation nationale des acquis scolaires (Vol 5)

UNESCO-UIS et ACER (2017). *Principles of Good Practice in Learning Assessment*. Montreal: UNESCO Institute for Statistics (UIS).

Wolff, Laurence (2007). *The Costs of Student Assessments in Latin America*. Washington, D.C.: PREAL.



Organisation
des Nations Unies
pour l'éducation,
la science et la culture



INSTITUT
DE STATISTIQUE
DE L'UNESCO



PARTENARIAT
MONDIAL *pour*
L'ÉDUCATION
*une éducation de qualité
pour tous les enfants*

@UNESCOstat
<http://uis.unesco.org>