

ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES POUR MESURER LES RÉSULTATS MONDIAUX D'APPRENTISSAGE BOÎTE À OUTILS

Relier les évaluations au cadre mondial des compétences

JANVIER 2023

UNESCO

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

La vocation première de l'UNESCO est de contribuer au maintien de la paix et de la sécurité dans le monde en resserrant la collaboration entre les nations par l'éducation, la science, la culture et la communication afin de promouvoir le respect universel de la justice, du droit, des droits de l'homme et des libertés fondamentales pour tous, sans distinction de race, de sexe, de langue ni de religion, que la Charte des Nations Unies reconnaît à tous les peuples.

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

L'Institut de statistique de l'UNESCO

L'Institut de statistique de l'UNESCO (ISU) est le bureau de statistique de l'UNESCO, chargé de rassembler, pour le compte des Nations-Unies, des statistiques mondiales dans les domaines de l'éducation, de la science et la technologie, et de la culture et la communication. L'ISU a été créé en 1999, avec la mission d'améliorer le programme de statistique de l'UNESCO ainsi que d'élaborer et de mettre à la disposition de la communauté internationale des statistiques actualisées, précises et pertinentes permettant d'orienter ses décisions dans un environnement politique, économique et social de plus en plus complexe et changeant.

Publié en 2023 par : l'Institut de statistique de l'UNESCO
C.P. 250 Succursale Centre-Ville
Montréal, Québec H3G 2K8 Canada

Courriel : uis.tcg@unesco.org
<http://www.uis.unesco.org>

© UNESCO-UIS 2023



Cette publication est disponible en Open Access sous la licence Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). En utilisant le contenu de cette publication, les utilisateurs acceptent d'être liés par les conditions d'utilisation du référentiel en libre accès de l'UNESCO (<http://www.unesco.org/open-access/terms-use-ccbysa-en>). Les désignations utilisées et la présentation des items tout au long de cette publication n'impliquent l'expression d'aucune opinion de la part de l'UNESCO concernant le statut juridique de tout pays, territoire, ville ou zone ou de ses autorités ou concernant la délimitation de ses frontières ou limites. Les idées et opinions exprimées dans cette publication sont celles des auteurs ; elles ne sont pas nécessairement celles de l'UNESCO et n'engagent pas l'Organisation.

Avant-propos

La méthodologie d'établissement de liens entre les politiques est un projet collaboratif de l'Institut de statistique de l'UNESCO (ISU). Le guide d'établissement de lien entre les politiques a été développé en partenariat avec le Groupe de la Banque mondiale et la Grande-Bretagne. Les idées et opinions exprimées dans cette publication sont celles des auteurs ; elles ne sont pas nécessairement celles des agences impliquées dans le développement de la boîte à outils.

TABLE DES MATIÈRES

TABLE DES MATIÈRES.....	iii
TABLEAUX	v
FIGURES.....	vii
ACRONYMES.....	ix
GLOSSAIRE DES TERMES.....	8
REMERCIEMENTS	10
CHAPITRE I. INTRODUCTION À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES.....	2
A. JUSTIFICATION DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES.....	2
B. PUBLIC VISÉ.....	4
C. PRÉSENTATION DU CADRE MONDIAL DES COMPÉTENCES	4
D. LE CADRE MONDIAL DE COMPÉTENCE ET LES NIVEAUX MINIMAUX DES COMPÉTENCES ...	8
E. PRÉSENTATION DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	10
F. ÉTAPES DANS L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES.....	11
G. UTILISATIONS ET AVANTAGES DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	13
H. UTILISER LE RÉFÉRENTIEL D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	13
CHAPITRE II. AUTO-ÉVALUATION DE LA PERTINENCE DE L'ÉVALUATION	16
A. COLLECTE DES PREUVES ET QUESTIONS À EXAMINER	16
B. CRITÈRES DE VALIDITÉ D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	16
C. ÉTAPES SUIVANTES.....	23
CHAPITRE III : LA MÉTHODE D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES.....	26
A. CHOIX DES ITEMS.....	26
B. FAMILIARISATION	27
C. TÂCHE 1 - ALIGNER L'ÉVALUATION SUR LE GPF.....	27
D. TÂCHE 2 – FAIRE CORRESPONDRE LES ITEMS D'ÉVALUATION AVEC LES GPL ET LES GPD.....	32
E. TÂCHE 3 : LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DE CRITÈRES DE RÉFÉRENCE	35
CHAPITRE IV. PRÉPARATION DE L'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	41
A. CHOISIR LES ANIMATEURS ET LES ANALYSTE DE L'ATELIER.....	41
B. PLANIFIER LA LOGISTIQUE DE L'ATELIER.....	42
C. SÉLECTIONNER ET INVITER LES PANÉLISTES DE L'ATELIER	44
D. DOCUMENTS ET ANALYSES	47
E. FORMER LES ANIMATEURS DE CONTENU.....	54
F. TEST TECHNIQUE	54
CHAPITRE V. MISE EN ŒUVRE DE L'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES	58
A. OUVERTURE.....	59
B. FAMILIARISATION	60
C. TÂCHE 1 - ALIGNEMENT	61
D. TÂCHE 2 : MISE EN CORRESPONDANCE.....	63
E. TÂCHE 3 - ANALYSE COMPARATIVE.....	65
F. ÉVALUATION	69
G. CONSEILS SUPPLÉMENTAIRES POUR ORGANISER DES ATELIERS À DISTANCE	70

CHAPITRE VI. AUTO-ÉVALUATION DES RÉSULTATS DE L'ATELIER	74
A. PRODUCTION DE LA DOCUMENTATION TECHNIQUE (UNE FOIS L'ATELIER TERMINÉ)	74
B. SOUMETTRE DES PREUVES À L'ISU	76
BIBLIOGRAPHIE	78
ANNEXE A – RESSOURCES CONNEXES	67
ANNEXE B – NIVEAU MINIMAL DE COMPÉTENCES MONDIAL	68
ANNEXE C : RÉSUMÉ DU MODÈLE D'AUTO-ÉVALUATION (PERTINENCE DE L'ÉVALUATION)	69
ANNEXE D – LISTE DE CONTRÔLE POUR LA PRÉPARATION DE L'ATELIER.....	71
ANNEXE E – PLANIFICATEUR D'ACTIVITÉS D'ATELIER.....	73
ANNEXE F – MODÈLE D'ESTIMATION DU BUDGET	76
ANNEXE G – DIAPOSITIVES DE FACILITATION DE L'ATELIER.....	77
ANNEXE H – FORMULAIRE D'ÉVALUATION DE L'ALIGNEMENT POUR LA TÂCHE I	109
ANNEXE I – FORMULAIRE D'ÉVALUATION D'ITEMS.....	110
ANNEXE J – STATISTIQUES DE PRÉCISION, D'EXACTITUDE ET DE COHÉRENCE	115
ANNEXE K – MODÈLE DE LETTRE D'INVITATION POUR LES OBSERVATEURS	118
ANNEXE L – MODÈLE DE LETTRE D'INVITATION POUR LES PANÉLISTES DES ATELIERS.....	120
ANNEXE M – INFORMATIONS DÉMOGRAPHIQUES DES PANÉLISTES.....	122
ANNEXE N – STATISTIQUES AVANT L'ATELIER.....	124
ANNEXE O – EXEMPLES DE DONNÉES DE RÉTROACTION ET INSTRUCTIONS	126
ANNEXE P – PROGRAMME DE L'ATELIER	128
ANNEXE Q – EXEMPLES D'ORDRE DU JOUR POUR UN ATELIER EN PRÉSENTIEL ET À DISTANCE	130
ANNEXE R – FORMULAIRE D'ÉVALUATION DE L'ATELIER.....	135
ANNEXE S – DIAPOSITIVES À L'INTENTION DES ANIMATEURS DE CONTENU	141
ANNEXE T – CALCULS DES CRITÈRES DE RÉFÉRENCE POUR L'ATELIER	142
ANNEXE U – MODÈLE DE CERTIFICAT D'APPRÉCIATION	145
ANNEXE V – RÉSUMÉ DU MODÈLE D'AUTO-ÉVALUATION (RÉSULTATS DE L'ATELIER)	146

TABLEAUX

Tableau 1 : exemple de mathématiques de 2 ^{ème} année tiré du GPF.....	6
Tableau 2 : Étapes dans l'établissement de liens entre les politiques.....	12
Tableau 3: Guides de navigation	14
Tableau 4 : Critères d'alignement des évaluations en mathématiques pour les années 1 à 9.....	19
Tableau 5 : Critères d'alignement des évaluations en lecture pour les années 1 à 9.....	19
Tableau 6 : Exemple de résumé des résultats d'alignement pour une évaluation de 3e année	31
Tableau 7 : Formulaire d'évaluation des items à utiliser avec la modification Angoff oui-non.....	38
Tableau 8 : Options pour organiser un atelier consacré à l'établissement de liens entre les politiques	42
Tableau 9 : brève description des séances de l'atelier.....	50
Tableau 10 : Objectif de la discussion, choses à faire et à ne pas faire par tâche	55
Tableau 11 : Résumé des tâches et des activités de l'atelier consacré à l'établissement de liens entre les politiques.....	59
Tableau 12 : informations requises pour rendre compte au regard de l'ODD 4.1.1	76
Table 13 : liste de contrôle pour la préparation de l'atelier	71
Tableau 14 : Planificateur d'activités d'atelier	73
Tableau 15 : modèle de formulaire d'évaluation de l'alignement.....	109
Tableau 16 : Exemple de formulaire d'évaluation d'items pour les évaluations non chronométrées	110
Tableau 17 : exemple de formulaire d'évaluation d'items pour les évaluations avec des questions à réponse construite.....	112
Tableau 18 : Exemple de formulaire d'évaluation d'items pour une évaluation de lecture chronométrée (en haoussa).....	113
Tableau 19 : Exemple de formulaire d'évaluation d'items pour les questions de compréhension de lecture conditionnelles (en haoussa).....	114
Tableau 20 : Tableau de distribution des données du modèle (CTT)	125
Tableau 21 : Tableau de distribution des données du modèle (IRT).....	125
Tableau 22 : tableau de données d'impact du modèle.....	127
Tableau 23 : Programme de l'atelier	128
Tableau 24 : Exemple d'ordre du jour pour un atelier en présentiel.....	130
Tableau 25 : Exemple d'ordre du jour pour la session de préparation à distance 1	132
Tableau 26 : Exemple d'ordre du jour pour la session de préparation à distance 2	132
Tableau 27 : Exemple d'ordre du jour pour la session 1 de l'atelier à distance	132
Tableau 28 : Exemple d'ordre du jour pour la session 2 de l'atelier à distance	133
Tableau 29 : exemple de programme pour la session 3 de l'atelier à distance.....	133
Tableau 30 : Exemple d'ordre du jour pour la session 4 de l'atelier à distance	133
Tableau 31 : Exemple d'ordre du jour pour la session 5 de l'atelier à distance	134
Tableau 32 : Exemple d'ordre du jour pour la session 6 de l'atelier à distance	134
Tableau 33 : Formulaire d'évaluation de la formation sur le GPF.....	135
Tableau 34 : Formulaire d'évaluation pour la formation à l'évaluation.....	136

Tableau 35 : Formulaire d'évaluation pour la tâche 1 – Alignement.....	136
Tableau 36 : formulaire d'évaluation pour la tâche 2 – Mise en correspondance.....	137
Tableau 37 : Formulaire d'évaluation pour la tâche 3 – Analyse comparative.....	138
Tableau 38 : Formulaire d'évaluation pour la tâche 3 – Série 2 d'analyse comparative.....	139

FIGURES

Figure 1 : Définir un ou trois critères de référence.....	8
Figure 2 : Exemple de critères de référence comparables pour diverses évaluations	10
Figure 3 : Exemple de points de référence comparables pour diverses évaluations	11
Figure 4 : Processus et avantages de l'établissement de liens entre les politiques.....	13
Figure 5 : Échelle d'alignement et nombre d'énoncés de connaissances et/ou de compétences auxquels un item s'aligne	28
Figure 6 : Exemple d'alignement d'un item avec le GPF et avec correspondance complète	29
Figure 7 : Exemple d'alignement d'un item sur le GPF avec correspondance partielle.....	29
Figure 8 : Exemple d'alignement d'un items sur le GPF sans correspondance	31
Figure 9 : Exemple de correspondance d'items avec les GPL et les GPD	34
Figure 10 : Items correspondants identifiés comme « sans correspondance » avec le GPF	34
Figure 11 : Processus d'évaluation des items pour la modification Angoff Oui-Non	36
Figure 12 : Niveau scolaire/complexité du texte des passages de lecture	37
Figure 13 : Activités de préparation à l'atelier consacré à l'établissement de liens entre les politiques.....	41
Figure 14 : composition des panélistes.....	45
Figure 15 : considérations relatives à la sécurité de l'évaluation.....	47
Figure 16 : Traduction du GPF.....	52
Figure 18 : Conseils aux animateurs sur la présentation du contexte.....	60
Figure 19 : Conseils aux animateurs sur la présentation du GPF	61
Figure 20 : Conseils aux animateurs sur la présentation de l'évaluation.....	61
Figure 21 : Conseils aux animateurs sur la présentation de l'alignement.....	62
Figure 22 : Conseils à l'intention des animateurs pour la tâche 1 – Aligner l'évaluation ou les évaluations sur le GPF.....	62
Figure 23 : Conseils à l'intention des animateurs pour l'examen des résultats de la tâche 1.....	63
Figure 24 : Conseils à l'intention des animateurs sur la présentation de la mise en correspondance de a tâche 2.....	63
Figure 25 : Conseils à l'intention des animateurs pour superviser l'activité de mise en correspondance de la tâche 2	64
Figure 26 : Conseils à l'intention des animateurs pour l'examen des résultats de la mise en correspondance de la tâche 2.....	64
Figure 27 : Conseils aux animateurs pour la présentation de l'analyse comparative mondiale	65
Figure 28 : Conseils aux animateurs pour la présentation de la méthode Angoff de la tâche 3.....	66
Figure 29 : Conseils aux animateurs sur la pratique Angoff de la tâche 3	66
Figure 30 : Conseils aux animateurs pour superviser la tâche 3 – Notations de la série 1	67
Figure 31 : Conseils aux animateurs pour partager les résultats de la série 1.....	67
Figure 32 : Conseils aux animateurs pour la présentation de la série 2 Angoff	68
Figure 33 : Conseils aux animateurs pour la présentation de la série 2 Angoff	68

Figure 34 : Conseils aux animateurs pour la présentation des résultats finaux.....	69
Figure 35 : Conseils aux animateurs pour la présentation du formulaire d'évaluation.....	69
Figure 36 : Conseils aux animateurs pour la clôture de l'atelier.....	69
Figure 37 : exemple de données normatives sur les notes des panélistes.....	126

ACRONYMES

ACER	Australian Council for Educational Research (Conseil australien pour la recherche en éducation)
DA	Dépasse au-delà les compétences minimales
CBA	Curriculum-Based Assessments (Évaluations basées sur le programme d'études)
RAC	Représentant de l'Agent Contractant
E3/ED	Bureau de la croissance économique, de l'éducation et de l'environnement
EGMA	Early Grade Math Assessment (Évaluation des mathématiques au primaire)
EGRA	Early Grade Reading Assessment (Évaluation de la lecture au primaire)
D_j	Dépasse les compétences minimales
FCDO	Foreign, Commonwealth and Development Office (Bureau des affaires étrangères, du Commonwealth et du développement)
GAML	Global Alliance to Monitor Learning (Alliance mondiale pour le suivi de l'apprentissage)
GPD	Global Proficiency Descriptor (Descripteur mondial de compétence)
GPF	Global Proficiency Framework (Cadre mondial de compétence)
GPL	Global Proficiency Level (Niveau mondial de compétence)
GRN	Global Reading Network (Réseau mondial de lecture)
ICAN	International Common Assessment of Numeracy (Evaluation commune internationale de la numératie)
JD	Juste dépasse les compétences minimales
JA	Juste Atteint les compétences minimales
JPA	Juste partiellement atteint les compétences minimales
A_j	Atteint les compétences minimales
MPL	Minimum Proficiency Level (Niveau minimal de compétence)
MSI	Management Systems International, Inc.
NAEP	National Assessment of Educational Progress (Évaluation des progrès dans le système éducatif)
NFER	National Foundation for Educational Research (Fondation nationale pour la recherche en éducation)
PAL	People's Action for Learning (Action populaire pour l'apprentissage)
PLT	Boîte à outils d'établissement de liens entre les politiques
PA_j	Partiellement atteint les compétences minimales
ODD	Objectif de Développement Durable 4.1.1
SE	Erreur standard (ou erreur type)
BEPH	Besoins éducatifs particuliers et handicaps
TEP Centre	Le Centre de partenariat pour l'éducation
ISU	L'Institut de statistique de l'UNESCO
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture
EM	Expert en la matière
USAID	U.S. Agency for International Development (Agence américaine pour le développement international)
USG	United States Government (Gouvernement des États-Unis)

GLOSSAIRE DES TERMES

La méthode Angoff – Une méthode de définition de normes dans laquelle les panélistes évaluent les items par GPL, puis font la moyenne des notes de tous les panélistes pour chaque GPL pour créer une référence.

Référence – Le score d'une évaluation qui indique que vous avez atteint un niveau de compétence.

Étendue de l'alignement – Couverture suffisante des domaines, des concepts et des sous-concepts du GPF par au moins un item d'évaluation.

Théorie classique des tests – Une théorie psychométrique basée sur l'idée que le score observé d'un individu à un test résulte de la somme d'une composante du score réel pour le candidat et d'une composante d'erreur aléatoire indépendante.

Normes de contenu – Ce que les apprenants doivent savoir et être capables de faire comme décrit dans le tableau GPF sur les connaissances et les compétences.

Ampleur de l'alignement – Couverture suffisante des items d'évaluation par le GPF.

Distracteur – Un ensemble de réponses plausibles mais incorrectes à l'item à choix multiples d'une évaluation.

Descripteur mondial de compétence (GPD) : une définition détaillée élaborée par des experts en la matière qui clarifie la part du contenu décrit dans les énoncés de connaissances et/ou de compétences du GPF qu'un apprenant devrait être capable de démontrer dans une matière à un niveau scolaire. Celles-ci sont parfois appelées normes de performance. Toutefois, les auteurs n'ont volontairement pas utilisé ce terme, car les pays ont leurs propres normes de performance qui peuvent différer des normes mondiales pour des raisons importantes. L'ensemble des GPD inclus dans le GPF n'est pas censé être de nature prescriptive mais plutôt pour faciliter la mesure au regard de l'ODD 4.1.1.

Niveau mondial de compétence (GPL) : les quatre niveaux de compétence ou de performance (ci-dessous répondent partiellement aux compétences minimales mondiales, satisfont partiellement aux compétences minimales mondiales, atteignent les compétences minimales mondiales et dépassent les compétences minimales mondiales) que les élèves peuvent atteindre pour tous les niveaux scolaires et matières ciblés. Le niveau de compétence minimale mondiale correspond à l'ODD 4.1.1, et les autres permettent aux pays de montrer les progrès réalisés pour que tous les étudiants atteignent ou dépassent ce niveau.

Données d'impact – Les données qui aident les panélistes à comprendre les conséquences de leurs jugements sur la population d'apprenants qui sont soumis à l'application des critères de référence recommandés par les panélistes.

Cohérence entre évaluateurs – Un indice qui indique l'accord ou le consensus global des panélistes parmi toutes les équipes possibles de panélistes.

Discrimination des items – La capacité d'un item à différencier les apprenants en fonction de leur compréhension du matériel testé, rapportée sur une échelle de -1 à +1.

Facilité d'item – La probabilité qu'un candidat réponde correctement à un item sur une échelle de 0 à 1.

Théorie de la réponse aux items – Un modèle mathématique de la relation fonctionnelle entre les performances sur un item de test, les caractéristiques de l’item de test et la position du candidat sur le concept mesuré.

Informations normatives – La répartition des références fixées par les panélistes, l’emplacement de chaque panéliste étant indiqué par une lettre de code ou un numéro connu d’eux seuls.

Normes de performance : quelle part du contenu décrit dans les énoncés de connaissances et/ou de compétences (normes de contenu) les apprenants sont censés être capables de démontrer. Voir également la définition du descripteur de compétence mondiale ci-dessus.

Établissement de liens entre les politiques pour mesurer les résultats d’apprentissage globaux : une méthode spécifique et non statistique qui utilise le jugement d’experts pour relier les scores des apprenants aux différentes évaluations aux niveaux des compétences minimales mondiales. L’établissement de liens entre les politiques comprend des processus d’alignement et de mise en correspondance entre les évaluations, le GPF et l’établissement de critères de référence.

Statistiques de difficulté des items – Informations sur la difficulté empirique des items (c-à-d. le pourcentage d’apprenants obtenant un item correct), qui donnent aux panélistes une idée approximative de la façon dont leurs jugements sur les items se comparent aux performances réelles de l’apprenant.

Erreur type (SE) – Une statistique qui indique l’erreur de mesure associée à un critère de référence (jugement du panéliste).

Énoncés de connaissances et/ou de compétences – Ce que les apprenants sont censés savoir et être capables de faire pour un niveau et un domaine, un concept et un sous-concept spécifiques. Les énoncés de connaissances et/ou de compétences sont parfois appelés normes de contenu. Toutefois, les auteurs n’ont volontairement pas utilisé ce terme, car les pays possèdent leurs propres normes de contenu qui peuvent différer des normes mondiales pour des raisons importantes. Les déclarations de connaissances et/ou de compétences incluses dans le GPF ne sont pas censées être de nature prescriptive mais plutôt pour faciliter la mesure au regard de l’ODD 4.1.1.

Liaison statistique – Méthodes qui utilisent des personnes ou des items communs pour relier les scores des apprenants aux différentes évaluations. Les méthodes de liaison statistique comprennent l’équation, l’étalonnage, la modération et la projection.

Énoncé – La question fait partie d’un élément à choix multiples lors d’une évaluation.

Méthode centrée sur les tests – Une famille de méthodes d’établissement de critères qui émettent des jugements basés sur un examen du matériel d’évaluation et des grilles de notation ; la méthode Angoff est incluse dans cette catégorie.

Évaluation chronométrée – Dans ce référentiel, le terme « évaluation chronométrée » est utilisé pour décrire une évaluation qui implique au moins un item/une tâche qui oblige les apprenants à entreprendre une activité dans un délai imparti, leur score étant lié à la quantité d’activités réalisées dans ce laps de temps. Par exemple, une évaluation de la maîtrise du langage verbal où les apprenants sont invités à lire autant de mots que possible en une minute.

Évaluation non chronométrée – Dans ce référentiel, le terme « évaluation non chronométrée » est utilisé pour décrire une évaluation dans laquelle il n’y a aucun item/tâche d’évaluation qui oblige les

apprenants à entreprendre une activité dans un délai imparti, leur score étant lié à la quantité d'activité qu'ils terminent à ce moment-là. Il s'agira souvent d'évaluations basées sur le programme (CBA) et elles peuvent avoir une limite de temps globale pour l'évaluation, mais pas pour des items individuels.

REMERCIEMENTS

Cette version originale du référentiel a été élaborée à la suite d'ateliers parrainés par le Bureau de l'éducation du Bureau pour la croissance économique, l'éducation et l'environnement (E3/ED) de l'Agence des États-Unis pour le développement international (USAID) et l'Institut de statistique de l'UNESCO (ISU). L'USAID et l'ISU de même que d'autres agences, dont le Groupe de la Banque mondiale (GBM), le Bureau britannique des affaires étrangères, du Commonwealth et du développement (FCDO) (anciennement le Département britannique pour le développement international) et la Fondation Bill & Melinda Gates, se sont montrés extrêmement favorables à l'introduction et à l'exploration des liens entre les politiques en tant que méthode de comparaison et d'agrégation des résultats des évaluations des apprenants dans et entre les pays.

L'équipe du projet tient à remercier Benjamin Sylla pour son leadership dans le développement du référentiel initial en tant que représentant de l'agent contractuel de l'USAID (RAC) du projet d'évaluation de la lecture et de l'accès, ainsi que le Dr Saima Malik, Rebecca Rhodes et le Dr Elena Walls de l'USAID E3/ED pour leur direction et leurs conseils tout au long du processus d'élaboration du projet initial. Silvia Montoya, directrice de l'ISU, a joué un rôle déterminant dans le soutien organisationnel. Jennifer Gerst du Réseau mondial de lecture (Global Reading Network, GRN) a joué un rôle clé dans l'animation des ateliers. Nous apprécions grandement toutes les contributions.

L'équipe du projet tient également à remercier les auteurs du référentiel initial : le Dr Abdullah Ferdous, Sean Kelly et le Dr Jeff Davis de Management Systems International (MSI), qui ont bénéficié du soutien de Melissa Chiappetta (un entrepreneur indépendant travaillant avec l'ISU, l'USAID et la Fondation Bill & Melinda Gates, qui a également apporté sa contribution en dirigeant le groupe de travail sur l'établissement de liens entre les politiques) ; Norma Evans d'Evans and Associates; Colin Watson du ministère de l'Éducation du Royaume-Uni ; et Carlos Fierros (NORC), Nathalie Liautaud et Ryan Aghabozorg (MSI).

Enfin, l'équipe tient à remercier l'ensemble des participants aux processus de développement, de pilotage et de révision du référentiel et du matériel, et en particulier, les ministères de l'Éducation et/ou l'agence ou le département spécialisé pour l'administration des évaluations nationales de l'apprentissage au Bangladesh, au Cambodge, à Djibouti, en Gambie, au Ghana, en Inde, au Kenya, au Lesotho, au Maroc, au Népal, au Nigéria, au Rwanda et au Sénégal. Ces projets pilotes ont été gérés par l'ISU, l'USAID et le GBM. Nous adressons également nos remerciements au réseau People's Action for Learning (PAL - Réseau d'action citoyenne pour l'apprentissage), The Education Partnership (TEP Centre) et Zizi Afrique, qui ont soutenu un projet pilote de Évaluation commune internationale de la numératie (International Common Assessment for Numeracy, ICAN). La participation mondiale aux activités de liaison politique a été conséquente, et nous espérons, qu'elle se poursuivra à l'avenir.

La deuxième phase de pilotage s'est déroulée en Inde, au Cambodge, au Népal, au Lesotho et en Zambie, techniquement soutenus par le CITO dans le cadre d'un contrat avec l'ISU et coordonnés par Melchior de Vries. Silvia Montoya et Shailendra Sigdel quant à eux, s'étaient chargés de sa coordination depuis l'ISU pour l'achever.

La version actuelle du référentiel a été mise à jour par l'ACER, sous contrat avec l'ISU, sur la base des commentaires du nombre de projets pilotes, des évaluations indépendantes menées par le CITO et la Fondation nationale pour la recherche en éducation (NFER), et des activités sur les enseignements tirés avec des groupes d'experts impliqués dans l'établissement de liens entre les politiques du monde entier. L'ACER a tout intérêt à soutenir l'utilisation d'une gamme d'outils permettant d'établir des références pour les évaluations nationales et internationales.

CHAPITRE I

CHAPITRE I. INTRODUCTION À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

A. JUSTIFICATION DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Alors que le nombre de pays participant à des évaluations des résultats d'apprentissage a connu une hausse significative au cours des deux dernières décennies, les méthodes permettant de comparer les résultats d'évaluation au sein des pays et entre les pays, ainsi que d'agréger ces résultats pour les rapports mondiaux, font défaut. Les ministères de l'Éducation, les responsables régionaux de l'évaluation, les donateurs internationaux de l'éducation, les partenaires et les autres parties prenantes ont besoin d'une méthode pour déterminer avec précision comment les résultats d'apprentissage se comparent entre les contextes d'un pays et d'un pays à l'autre, et comment les pays et les donateurs peuvent rendre compte des progrès dans des domaines clés tels que la lecture et les mathématiques. Ces informations sont essentielles afin d'identifier les lacunes dans les résultats d'apprentissage en vue de concentrer les ressources sur les domaines et les populations qui en ont le plus besoin.

Le principal défi lié à la réalisation de comparaisons mondiales et d'agrégations de résultats d'évaluation est que les pays utilisent généralement différents outils d'évaluation avec des niveaux de difficulté variables. Relier les différentes évaluations à une échelle commune permet de résoudre ce problème. Le lien peut être établi soit de manière statistique, en ayant recours à des items communs entre les évaluations ou en demandant aux apprenants communs de passer plus d'une évaluation, soit de manière non statistique, en se basant sur des jugements d'experts. Bien que les méthodes statistiques soient souvent associées à des niveaux de précision plus élevés, elles ne sont pas toujours dans la pratique ou financièrement réalisables et impliquent plusieurs prérequis méthodologiques.

En conséquence, ce référentiel décrit une méthode de jugement non statistique appelée établissement des liens entre les politiques pour mesurer les résultats d'apprentissage mondiaux (lien politique en abrégé), qui a également été appelée modération sociale.¹ L'Institut de statistique de l'UNESCO (ISU) a inclus l'établissement des liens entre les politiques dans sa liste de méthodologies acceptables pour l'établissement de rapports au regard de l'objectif de développement durable (ODD) 4.1.1 :

Proportion d'enfants et de jeunes : (a) en classes de 2/3e année primaire ; (b) à la fin de la primaire ; et (c) à la fin du secondaire inférieur atteignant au moins un niveau de compétence minimum en (i) lecture et (ii) mathématiques, par sexe.

D'autres organisations donatrices – notamment l'USAID, le FCDO, le Groupe de la Banque mondiale, la Fondation Bill & Melinda Gates et l'UNICEF – ont manifesté leur intérêt pour l'utilisation ou le soutien à l'établissement de liens entre les politiques pour établir des critères de référence pour les évaluations nationales et internationales, ce qui faciliterait l'établissement de rapports sur indicateurs mondiaux clés

¹ L'approche d'établissement de liens entre les politiques a été proposée en septembre 2017 lors d'une réunion de l'Alliance mondiale pour le suivi de l'apprentissage (GAML), puis de nouveau en août 2018 lors d'un atelier mondial organisé par l'USAID. En février 2019, l'USAID a publié un document sur les liens entre les politiques, avec le soutien technique de Management Systems International (MSI). Un groupe de 30 experts internationaux en la matière (EM) a produit le premier cadre mondial des compétences (GPF) en avril et mai 2019 couvrant les niveaux 2 à 6. La première version du référentiel pour l'établissement de liens entre les politiques a été produite en septembre 2019 pour guider les pilotes. Une autre version du GPF a été produite par un groupe élargi de EM en octobre 2020, parallèlement à une version révisée du référentiel. La deuxième ébauche du GPF a ajouté la 1ère et la 7ème à la 9ème année. La version actuelle du référentiel a été produite en décembre 2022.

liés à la lecture et aux mathématiques et permettrait également aux pays de fixer des objectifs d'apprentissage pour l'amélioration à long terme des résultats d'apprentissage.^{2,3} Avec l'ISU, ces agences ont formé un groupe de travail pour développer la méthode d'établissement de liens entre les politiques. Une version antérieure de ce référentiel a été utilisée pour tester la méthode d'établissement de liens entre les politiques dans trois pays d'octobre 2019 à mars 2020, après quoi elle a été révisée – avec les contributions du groupe de travail et d'un organisme d'évaluation indépendant (la Fondation nationale pour la recherche en éducation [NFER]) – pour cette version actuelle. La version actuelle du référentiel a été mise à jour à la suite d'autres projets pilotes menés dans cinq pays entre mars 2021 et mai 2022 et des recommandations de l'évaluation indépendante. L'évaluation NFER de la méthode, financée par la Fondation Bill & Melinda Gates, est en cours et continuera à éclairer les futurs changements apportés à la méthode.

Ce référentiel a été conçu pour établir des liens entre les politiques à l'aide du cadre mondial des compétences (Global Proficiency Framework (GPF)) (disponible sur Edulinks et sur le site internet de l'ISU), qui est décrit en détail ci-dessous. Le GPF est composé d'attentes internationalement convenues concernant les connaissances et/ou les compétences que les apprenants ayant les compétences minimales devraient posséder (ces énoncés de connaissances et/ou de compétences sont parfois appelées normes de contenu) et la mesure dans laquelle ils devraient être capables de les démontrer (appelés dans le GPF descripteurs de compétences mondiales, parfois appelés normes de performance) qui forment une échelle commune pour les rapports mondiaux sur les résultats des apprenants en lecture et en mathématiques de la 1^{re} à la 9^e année^{4,5}. Cependant, même si le référentiel a été développé pour aider les pays et les organisations d'évaluation régionales et internationales à établir des références pour les rapports mondiaux, elle peut également être utilisée pour définir des références nationales pour les rapports nationaux sur les évaluations existantes. Le gouvernement d'un pays peut choisir de fixer des références nationales et mondiales pour la même évaluation, et ces références pourraient être les mêmes si les cadres nationaux sont alignés sur le GPF et si les références sont définies en utilisant la même approche. Cependant, certains pays peuvent choisir de maintenir leurs propres normes nationales, distinctes des normes mondiales décrites dans le GPF. Les pays peuvent le faire pour des raisons telles que le choix d'enseigner les connaissances et les compétences à des niveaux scolaires différents de ceux représentés dans le GPF ou parce qu'ils souhaitent que leurs normes nationales intègrent des connaissances et des compétences supplémentaires non prises en compte dans le GPF. Dans de tels cas, les pays pourraient choisir de fixer des critères distincts pour les rapports nationaux et les rapports mondiaux.

² La Fondation Bill & Melinda Gates a commandé une évaluation en 2019 visant à évaluer de manière empirique l'acceptabilité de l'établissement de liens entre les politiques en tant que méthode permettant de lier les résultats d'évaluation à l'ODD 4.1.1. Les recommandations issues des observations pilotes de l'évaluation ont été intégrées à cette révision, bien que l'évaluation complète, notamment une étude de validation, reste en suspens. L'adhésion de la fondation à la méthode est conditionnée aux résultats de cette évaluation.

³ Un critère de référence est un seuil numérique sur une évaluation qui indique qu'un apprenant a atteint un niveau de compétence.

⁴ Les auteurs ont intentionnellement omis d'utiliser le terme « normes de contenu » dans le GPF car les pays possèdent leurs propres normes de contenu qui peuvent différer des normes mondiales pour des raisons majeures. Les déclarations de connaissances et/ou de compétences incluses dans le GPF ne sont pas censées être de nature normative mais visent plutôt à faciliter la mesure au regard de l'ODD 4.1.1.

⁵ Les auteurs ont intentionnellement omis d'utiliser le terme « normes de performance » dans le GPF car les pays possèdent leurs propres normes de performance qui peuvent différer des normes mondiales pour des raisons importantes. L'ensemble des GPD inclus dans le GPF n'est pas censé être de nature normative, mais vise plutôt à faciliter la mesure au regard de l'ODD 4.1.1.

Ce référentiel définit les normes essentielles pour la qualité de l'évaluation à utiliser pour l'établissement de liens entre les politiques et les résultats de la méthode d'établissement de liens entre les politiques. Si ces normes ne sont pas respectées, les résultats ne seront pas acceptés pour les rapports au regard de l'ODD 4.1.1. Lorsqu'elles ne sont pas respectées (par exemple, lorsque l'évaluation n'est pas suffisamment alignée sur le GPF ou que le plan d'échantillonnage ne produit pas de résultats suffisamment représentatifs de la population souhaitée), il peut toujours y avoir une justification pour la mise en œuvre de la méthode de liaison des politiques. Ce référentiel indiquera clairement les avantages de la poursuite de certaines activités dans une perspective de renforcement des capacités.

B. PUBLIC VISÉ

Ce référentiel a été créé pour être utilisé par les gouvernements nationaux et les agences d'évaluation (pour les évaluations multinationales) et leurs partenaires. Tous ceux qui l'utilisent, y compris les agences d'évaluation, doivent se coordonner étroitement avec le(s) gouvernement(s) national (aux) concerné(s), car ce sont les gouvernements qui, en fin de compte, rendront compte des résultats au regard de l'ODD 4.1.1.

C. PRÉSENTATION DU CADRE MONDIAL DES COMPÉTENCES

Le GPF a été créé pour répondre à l'appel lancé par le Rapport mondial de suivi sur l'éducation (Global Education Monitoring Report, GEMR), chargé de suivre les progrès vers l'atteinte de l'ODD 4, pour créer « des définitions communes de ce que sont les « résultats d'apprentissage pertinents et efficaces » afin qu'ils puissent être comparés dans tous les pays et surveillés à l'échelle mondiale. La méthode d'établissement de liens entre les politiques décrite dans ce référentiel nécessite cet ensemble commun de descripteurs de compétences mondiaux (parfois appelés normes de performance) par niveau scolaire et domaine auquel les pays peuvent lier leurs évaluations pour produire des indicateurs et faire rapport à l'ISU. En utilisant une approche d'analyse comparative standardisée, les résultats de différents pays et les évaluations liées aux normes GPF pour leur niveau et leur matière peuvent ensuite être comparés, regroupés et suivis. Par exemple, toutes les évaluations de fin de lecture du primaire peuvent être liées au GPF de lecture de 5e année, ce qui permet ensuite aux pays de produire les indicateurs de l'ODD 4.1.1(b) basés sur les descripteurs de compétences pour comparer, agréger et suivre les résultats de ces évaluations en lecture de fin de cycle du primaire.

Alors que les pays définissent quelles connaissances et/ou compétences les apprenants doivent acquérir dans quelles classes en fonction de leurs contextes individuels et articulent ces informations par le biais de normes, de programmes et d'évaluations nationaux, le GPF définit les connaissances et les compétences importantes que tous les enfants et les jeunes doivent atteindre, peu importe où ils vivent dans le monde.

Une équipe de plus de 60 experts en matière de lecture et de mathématiques (EM) du monde entier, ayant tous une expérience de travail dans plusieurs pays et contextes, s'est réunie pour créer le GPF. Le GPF définit, pour la lecture et les mathématiques à l'école primaire, le niveau de compétence minimale mondiale que les apprenants sont censés démontrer à la fin de chaque année (de la première à la neuvième année). Les EM sont parvenus à un consensus sur les énoncés de connaissances et/ou de compétences (parfois appelés normes de contenu) et sur les descripteurs de performance mondiale (GPD) (parfois appelés normes de performance) décrits dans le GPF sur la base de leur connaissance des progressions de développement et du cadre mondial de contenu (GPF) de l'ISU, basé sur 73 programmes d'études et cadres d'évaluation de 25 pays pour la lecture et 115 cadres d'évaluation de 53 pays pour les

mathématiques.^{6,7} Le GPF a également été examiné et documenté par des experts en langues et des personnes possédant une expertise dans le travail avec les étudiants handicapés et dans l'éducation dans les zones touchées par des crises et des conflits. Il était important que le GPF soit fondé sur le cadre de contenu et l'expérience d'experts dans divers contextes pour garantir que les normes décrites dans le document soient alignées sur les normes de contenu et les programmes d'études nationaux existants et n'imposent pas d'attentes plus élevées aux apprenants.

Un exemple tiré d'une partie du GPF en mathématiques de troisième année est présenté dans le **tableau I**. Il contient les domaines, les concepts, les sous-concepts, les énoncés de connaissances et/ou de compétences, ainsi que les GPD pour les trois premières catégories de performance sur quatre, appelés niveaux mondiaux de compétence (GPL). Notez que la catégorie de performance la plus basse, « En dessous du niveau de compétence minimale mondiale », n'a pas besoin de GPD puisqu'elle inclut tous les apprenants qui ne répondent pas aux attentes décrites dans le niveau « Satisfait partiellement au niveau de compétence minimale mondiale ».

⁶ Voir la note de bas de page précédente pour une chronologie du développement du GPF.s

⁷ Voir UNESCO (2018a, 2018b) dans les références de ses cadres de contenu global pour la lecture et les mathématiques. Notez que ces cadres ne sont pas classés par niveau scolaire et n'ont pas de descripteurs par niveau mondial de compétence (GPL).

Tableau 1 : exemple de mathématiques de 2^{ème} année tiré du GPF

Domaine	Concept	Sous-concept	Connaissances ou compétences (normes de contenu)	Niveaux de compétences minimales mondiales et descripteurs (normes de performance)		
				Partiellement atteint les compétences minimales mondiales	Atteint les compétences minimales mondiales	Dépasse les compétences minimales mondiales
Nombre et opérations	Nombres entiers	Identifier et compter en nombres entiers, et repérer leur grandeur relative	Compter, lire et écrire des nombres entiers	Compter en nombres entiers jusqu'à 30.	Compter en nombres entiers jusqu'à 100.	Compter à l'envers à partir de 20.
				Lire et écrire des nombres entiers jusqu'à 30 en lettres et en chiffres.	Lire et écrire des nombres entiers jusqu'à 100 en lettres et en chiffres.	N/D
			Comparer et classer des nombres entiers	Comparer et classer des nombres entiers jusqu'à 30.	Comparer et classer des nombres entiers jusqu'à 100.	N/D
			Compter par sauts en avant et en arrière	N/D	Compter par sauts de deux ou de dizaines.	Compter par sauts en arrière de dizaines.
		Représenter des nombres entiers de manière équivalente	Déterminer ou identifier l'équivalence entre des nombres entiers représentés sous forme d'objets, d'images et de chiffres	Identifier et représenter l'équivalence entre des quantités entières jusqu'à 10 représentées sous forme d'objets, d'images et de chiffres (<i>p. ex., lorsqu'on vous donne une image de 10 objets et d'autres images de différents nombres d'objets, sélectionner l'image qui possède le même nombre d'objets ; associer un chiffre avec le nombre approprié d'objets.</i>)	Identifier et représenter l'équivalence entre des quantités entières jusqu'à 30 représentées sous forme d'objets, d'images et de chiffres (<i>p. ex., lorsqu'on vous donne une image de 30 fleurs, identifiez l'image qui contient le nombre de papillons qui serait nécessaire pour que chaque fleur ait un papillon ; étant donné une image de 19 formes, dessiner 19 formes supplémentaires.</i>)	Utiliser des concepts de valeur de position pour les dizaines et les unités (<i>p. ex., composer ou décomposer un nombre entier à deux chiffres en utilisant une phrase numérique telle que $35 = 3$ dizaines et 5 unités, $35 = 30 + 5$, ou en utilisant des liaisons numériques ; déterminer la valeur d'un chiffre à la place des dizaines et des unités.</i>)

Comme le montre le **tableau I**, afin de définir le contenu de chaque niveau et matière, le GPF est organisé hiérarchiquement, c.-à-d. du général au spécifique, avec des domaines, des concepts et des sous-concepts. Les énoncés de connaissances et/ou de compétences associés aux sous-concepts démontrent ce que les apprenants doivent savoir et être capables de faire par niveau et par matière.

En plus des contrats de sous-traitance, il existe les GPD, qui décrivent la part du contenu des connaissances et des compétences que les apprenants doivent démontrer pour être considérés comme ayant une compétence minimale. Chacun des GPL est caractérisé par une définition – appelée définition politique – qui s'applique à tous les niveaux et à toutes les matières. Les quatre définitions – pour les quatre catégories de performance, ou GPL – sont fournies ci-dessous et également incluses à l'**annexe B** :

- **En dessous du niveau partiellement atteint les compétences minimales mondiales** : les apprenants ne possèdent pas les connaissances et les compétences de base correspondant à leur niveau. En conséquence, ils ne peuvent pas accomplir les tâches les plus élémentaires correspondant à leur niveau de classe.
- **Partiellement atteint les compétences minimales mondiales** : les apprenants ont des connaissances et des compétences partielles pour leur niveau. En conséquence, ils peuvent partiellement accomplir des tâches de base correspondant à leur niveau de classe.
- **Atteint les compétences minimales mondiales** : les apprenants ont des connaissances et des compétences suffisantes pour leur niveau. En conséquence, ils peuvent accomplir avec succès les tâches de base adaptées à leur niveau de classe.

- **Dépasse les compétences minimales mondiales** les apprenants ont des connaissances et des compétences suffisantes pour leur niveau. En conséquence, ils peuvent accomplir avec succès les tâches de base adaptées à leur niveau de classe.

Le groupe de travail sur l'établissement de liens entre les politiques a élaboré les quatre niveaux grâce à une consultation approfondie avec les parties prenantes nationales et internationales. Ils sont destinés à permettre aux pays de suivre et de rendre compte des progrès au fil du temps, dans le but qu'un pourcentage croissant d'apprenants passe de la catégorie « en dessous du niveau partiellement atteint les compétences minimales mondiales » à la catégorie « Partiellement atteint les compétences minimales mondiales » et éventuellement à la catégorie « Atteint les compétences minimales mondiales » voire « Dépasse les compétences minimales mondiales ».

Les GPD définissent ce qui est attendu des apprenants dans les trois derniers GPL (il n'est pas nécessaire d'avoir des GPD pour le niveau « Partiellement atteint les compétences minimales mondiales », car tous les apprenants qui ne satisfont pas au niveau « Partiellement atteint les compétences minimales mondiales » entreront dans cette catégorie) pour les élèves de la première à la neuvième année en lecture et en mathématiques. Ils décrivent le contenu que les apprenants doivent connaître et être capables de réaliser par rapport aux énoncés de connaissances et/ou de compétences requises par niveau et par matière. Par exemple, en lecture, le GPF indique qu'un apprenant qui atteint le niveau de compétence minimale mondiale en troisième année devrait être capable d'identifier le sujet général dans un texte continu de troisième année lorsque le sujet est important mais n'est pas explicitement énoncé. En mathématiques, un apprenant qui atteint le niveau de compétence minimale mondiale en troisième année devrait être capable de comparer et d'ordonner des nombres entiers jusqu'à 1 000.

Notez que l'établissement de liens entre les politiques est conçu pour être utilisé avec les quatre GPL. Cela fournit des informations pour rendre compte de certains indicateurs de donateurs, tels que les indicateurs d'aide étrangère (« F ») de l'USAID⁸. Cependant, le gouvernement d'un pays ou d'une agence d'évaluation peut choisir d'utiliser uniquement le niveau Atteint le GPL appropriée, ce qui est suffisant pour établir des rapports sur l'ODD 4.1.1 (voir la **figure I** pour certains critères que les pays et les organismes d'évaluation peuvent les prendre en compte lorsqu'ils décident du nombre de critères qu'ils doivent définir). Cependant, comme mentionné, l'établissement de critères de référence pour les trois niveaux supérieurs est encouragé, car cela permettra aux pays et aux partenaires de mieux démontrer les progrès réalisés au fil du temps vers la satisfaction des exigences de l'ODD 4.1.1, même si cela nécessite une évaluation suffisamment longue (voir le critère d'alignement au **chapitre II**). Les pays ou partenaires communiquant sur les indicateurs de l'USAID devront établir des références pour les trois niveaux de performance les plus élevés, puisque certains des indicateurs « F » mesurent l'amélioration d'un niveau de performance à un autre.

⁸ Les indicateurs « F » de l'USAID sont formulés comme suit : « les apprenants atteignent un niveau de compétence minimale en lecture/mathématiques » et sont donc liés aux normes du GPF.

Figure I : Définir un ou trois critères de référence

Trois critères de référence sont recommandés car :

- Ils facilitent davantage le suivi des progrès vers la réalisation des objectifs de l'ODD 4.1.1
- Ils sont conformes aux exigences en matière de rapports en comparaison avec les indicateurs d'aide étrangère de l'USAID
- Ils permettent aux pays de mieux déceler les lacunes en matière d'apprentissage et de cibler ceux qui en ont le plus besoin

Cependant, un seul critère de référence est nécessaire pour établir des rapports sur l'ODD 4.1.1. Il peut être judicieux pour les pays/agences d'évaluation de définir un critère de référence si :

- leurs évaluations sont courtes et il est peu probable qu'elles aient une gamme de scores suffisamment large pour faciliter plusieurs critères de référence uniques
- Ils n'ont pas de partenariat avec l'USAID
- Ils ont d'autres normes d'évaluation nationales pour lesquelles ils souhaitent également établir des critères de référence pour le suivi des besoins avec leur pays

D. LE CADRE MONDIAL DE COMPÉTENCE ET LES NIVEAUX MINIMAUX DES COMPÉTENCES

L'ISU a élaboré des niveaux de compétences minimales (MPL) pour soutenir la notification de l'ODD 4.1.1. Ceux-ci décrivent les connaissances de base dans un domaine, tels que mesurés par des évaluations d'apprentissage, à a) la fin du premier cycle du primaire (2^{ème}/3^{ème} année), b) la fin du primaire et c) la fin du premier cycle du secondaire. Ceux-ci ont été utilisés dans le développement du GPF, bien que les MPL soient définis en termes de stade de scolarité et que le GPF soit défini en termes de niveau d'un apprenant, il existe quelques différences.

La fin des MPL du premier cycle du primaire (appelées « 2^{ème}/3^{ème} année » au point 4.1.1) est décrite en termes d'un standard unique pour la lecture et d'un standard unique pour les mathématiques. L'alignement avec le GPF pour la lecture et les mathématiques est le plus proche des descriptions « Atteint les compétences minimales mondiales » pour la 2^{ème} année.

Les MPL de fin de primaire (4.1.1b) sont également décrits en termes d'un seul standard pour la lecture et d'un seul standard pour les mathématiques. L'alignement avec le GPF pour la lecture et les mathématiques est le plus proche des descriptions « Atteint les compétences minimales mondiales » pour la 5^e année.

Enfin, les MPL de fin de premier cycle du secondaire (4.1.1c) sont décrits en termes d'un seul standard pour la lecture et d'un seul standard pour les mathématiques. L'alignement avec le GPF pour la lecture et les mathématiques est le plus proche des descriptions « Atteint les compétences minimales mondiales » pour la 8^e année.

Pour rendre compte de l'ODD 4.1.1 (a), (b) et (c), les pays devront utiliser le GPF aux niveaux 2, 5 et 8 respectivement pour établir un lien entre les politiques, quel que soit le niveau des apprenants qui passent l'évaluation. Il s'agit de garantir la comparabilité entre les pays dans la norme attendue pour les rapports sur l'ODD 4.1.1. Les pays peuvent choisir de mettre en œuvre une politique liée à la note du GPF qui correspond à leur évaluation (par exemple, en utilisant le GPF en 4^e année pour une évaluation

administrée aux apprenants de 4e année), mais cela ne serait pas accepté par l'ISU pour le rapport sur l'ODD 4.1.1.

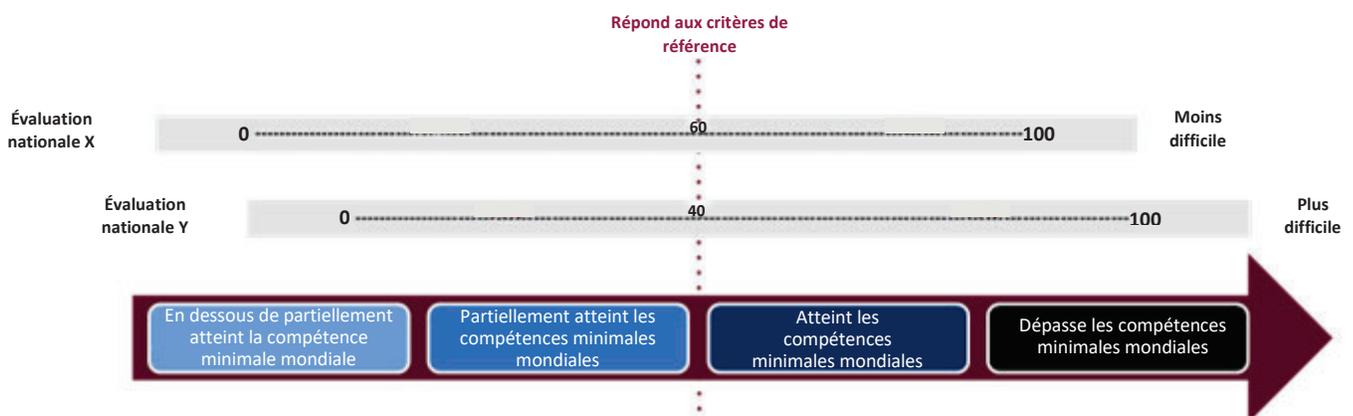
E. PRÉSENTATION DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

L'établissement de liens entre les politiques est une méthode qui permet aux pays/agences d'évaluation de lier leurs évaluations à l'ODD 4.1.1 et de déterminer les critères de référence de ces évaluations pour atteindre la compétence minimale mondiale.⁹ Il rassemble des animateurs principaux, des animateurs de contenu, des panélistes et des observateurs de représentants du gouvernement et d'agences d'évaluation pour mener à bien ce processus. Les rôles et qualifications de chacun de ces groupes sont présentés au chapitre III.

Étant donné que le GPF est utilisé comme un critère de référence – ou critère commun – pour établir des liens entre les politiques, ces critères représentent le même niveau de performance pour ces différentes évaluations telles que définies par les GPD, quelle que soit la difficulté ou la langue des évaluations, même si les critères sont fixés à différents endroits (scores numériques) sur les différentes évaluations (sauf si les évaluations sont de difficulté équivalente).

Par exemple, comme le montre la **figure 2**, deux évaluations différentes auront très probablement des critères différents pour atteindre la compétence minimale mondiale en raison de la difficulté inégale de ces évaluations. Pour un niveau et une matière donnés, les évaluations les moins difficiles auront des critères plus élevés et les évaluations plus difficiles auront des critères plus bas. Dans cet exemple, le pays X et le pays Y ont des évaluations nationales avec des échelles de 0 (minimum) à 100 (maximum) points. Ils relient leurs évaluations au GPF. L'évaluation nationale X – qui est moins difficile – a un critère de compétence minimale mondiale de 60 points, tandis que l'évaluation nationale Y – qui est plus difficile – a un critère de compétence minimale mondiale de 40 points. En théorie, un apprenant dont le niveau de compétence atteint tout juste le niveau de compétence minimale mondiale et qui passe les deux évaluations obtiendra 60 points pour l'évaluation la moins difficile et 40 points pour l'évaluation la plus difficile. Comme le montre le schéma ci-dessous, les évaluations varient en difficulté, mais le barème commun du GPF reste constant, les référentiels liés au GPF sont donc équivalents.

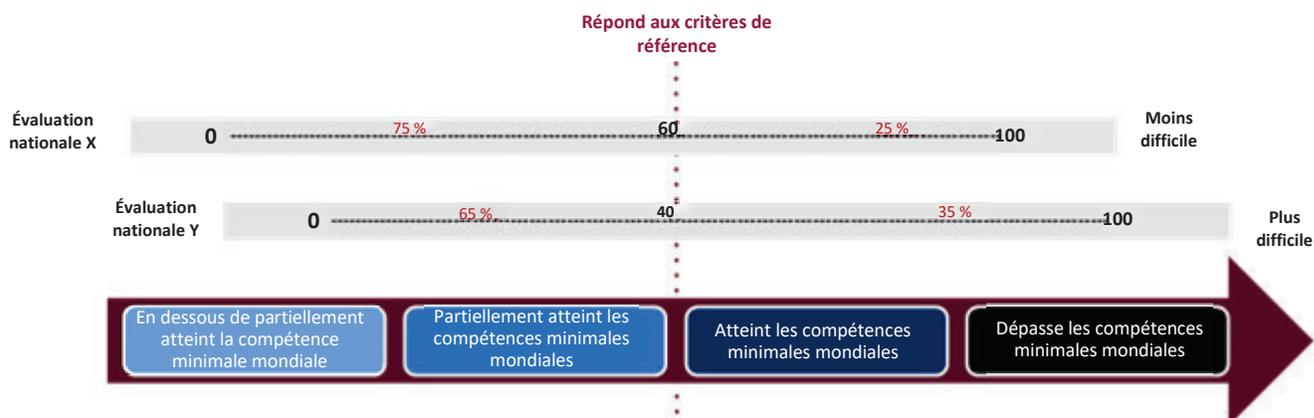
Figure 2 : Exemple de critères de référence comparables pour diverses évaluations



⁹ Les critères de référence d'une évaluation déterminent si un apprenant est classé dans une catégorie ou un niveau de performance ; ils sont également connus sous le nom de scores de coupure, points de coupure, seuils ou limites.

En fixant les critères de référence pour différentes évaluations basées sur les mêmes descripteurs du GPF, les évaluations sont liées par leurs critères équivalents, c'est-à-dire les critères de chaque évaluation qui correspondent à l'atteinte des compétences minimales mondiales. Dans cet exemple, comme le montre la **figure 3**, le processus a déterminé que 25 % des étudiants du pays X et 35 % des étudiants du pays Y ont atteint le niveau de compétence minimale.

Figure 3 : Exemple de points de référence comparables pour diverses évaluations



Pour établir les critères de référence, l'établissement de liens entre les politiques utilise une procédure d'analyse comparative Angoff internationalement reconnue, standardisée, centrée sur les tests et révisée. La procédure Angoff demande à des groupes de EM nationaux, appelés panélistes, de porter un jugement sur les items utilisés dans les évaluations nationales. Les panélistes comprennent des maîtres enseignants et des experts en programmes d'études du pays (des pays dans le cas des évaluations multinationales) qui comprennent les performances des apprenants pour des niveaux et des matières spécifiques. Ils suivent la procédure Angoff pour 1) examiner le ou les instruments d'évaluation du pays/de l'agence d'évaluation par rapport aux GPD et 2) estimer les performances des apprenants de chacune des catégories GPL à l'évaluation. Les planificateurs et les animateurs organisent et animent des ateliers séparés par niveau, matière et langue avec différents groupes de panélistes afin de définir les critères équivalents pour ces évaluations.

F. ÉTAPES DANS L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

L'établissement de liens entre les politiques visant à mesurer les résultats d'apprentissage mondiaux comporte six étapes qui doivent être réalisées pour faciliter l'établissement de rapports mondiaux, comme le montre le **tableau 2**. Les pays/agences d'évaluation et leurs partenaires doivent franchir chacune de ces étapes pour que leurs résultats soient acceptés pour présenter un rapport en fonction de l'ODD 4.1.1 et aux indicateurs « F » de l'USAID. Ce référentiel couvre les étapes 2, 3, 4 et 5. Le **tableau 2** fournit des informations sur les ressources disponibles pour soutenir les autres étapes.

Tableau 2 : Étapes dans l'établissement de liens entre les politiques

N°	Étapes dans l'établissement de liens entre les politiques	Objectif	Rôles/Responsabilités	Ressources (disponibles sur le site internet de l'ISU)
1	Engagement initial	Pour les pays (ou les agences d'évaluation en coordination avec les gouvernements nationaux concernés) afin de déterminer si, pour une évaluation spécifique, l'établissement de liens entre les politiques est l'approche préférée, soit au niveau national, soit au niveau régional/étatique, pour établir un rapport au regard de l'ODD 4.1.1	Les gouvernements nationaux/agences d'évaluation peuvent effectuer cette étape eux-mêmes ou ils peuvent demander/recevoir le soutien de leurs partenaires tels que, l'ISU, les donateurs et/ou les sous-traitants chargés l'établissement de liens entre les politiques. Il est essentiel que les gouvernements nationaux s'approprient ce processus et soient disposés à fournir les informations, rapports et données nécessaires à toutes les personnes concernées au moment opportun pour appuyer le travail entrepris. L'appropriation du processus par les gouvernements des pays soutiendra également le développement des capacités, dans le but souhaité qu'ils soient en mesure de gérer eux-mêmes les ateliers à venir.	<ul style="list-style-type: none"> • Rapport sur les acquis d'apprentissage dans l'éducation de base : options du pays pour l'indicateur 4.1.1¹⁰
2	Auto-évaluation de la pertinence de l'évaluation pour l'établissement de liens entre les politiques	Déterminer si la fiabilité, la validité et l'alignement de l'évaluation avec le GPF répondent aux exigences pour procéder à l'établissement de liens entre les politiques en vue de la rédaction de rapports mondiaux ; et déterminer le nombre de critères à fixer sur l'évaluation en fonction de sa durée. (Lorsqu'une évaluation n'est pas jugée appropriée pour l'établissement d'un rapport mondial, des activités seront proposées pour le renforcement des capacités)	Gouvernements nationaux/agences d'évaluation avec/sans le concours des partenaires et observateur indépendant approuvé par l'ISU	<ul style="list-style-type: none"> • Référentiel d'établissement de liens entre les politiques (Chapitre II) • Modèle de rapport d'auto-évaluation (Annexe C)
3	Préparation d'un atelier d'établissement de liens entre les politiques	Identifier/confirmer les animateurs (si ce n'est pas fait), inviter les panélistes, préparer le matériel et réserver un lieu	Gouvernements nationaux/agences d'évaluation avec/sans le concours des partenaires	<ul style="list-style-type: none"> • Référentiel d'établissement de liens entre les politiques (Chapitre IV) • Check-list pour la préparation de l'atelier (Annexe D)
4	Mise en œuvre de l'atelier de l'établissement de liens entre les politiques (composé de trois étapes - alignement, correspondance et analyse comparative)	Établir des points de référence et documenter les détails concernant le processus suivi	Gouvernements nationaux/agences d'évaluation avec/sans le concours des partenaires	<ul style="list-style-type: none"> • Référentiel d'établissement de liens entre les politiques (Chapitre V)
5	Auto-évaluation des résultats de l'atelier d'établissement de liens entre les politiques	Déterminer si la fiabilité et la validité de l'atelier répondent aux critères d'établissement de rapports mondiaux	Gouvernements nationaux/agences d'évaluation avec/sans le concours des partenaires et observateur indépendant approuvé par l'ISU	<ul style="list-style-type: none"> • Référentiel d'établissement de liens entre les politiques (Chapitre VI) • Modèle de rapport d'auto-évaluation (Annexe V)
6	Communiquer les résultats pour l'ODD 4.1.1 (et/ou d'autres indicateurs des donateurs)	Pour qu'un pays soit pris en compte dans un rapport mondial	Gouvernements nationaux avec/sans le concours des partenaires	<ul style="list-style-type: none"> • Protocole de reporting sur l'indicateur mondial 4.1.1 des ODD (mars 2022)¹¹ • Directives individuelles des donateurs

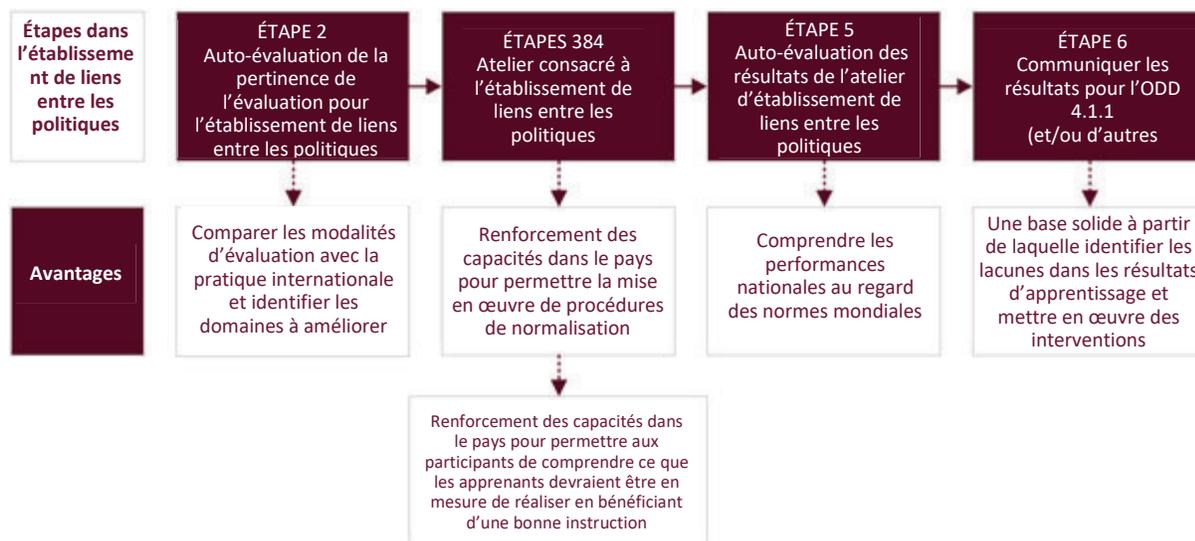
¹⁰ https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2022/08/Countrys-reporting-option-_Zambia_AAEA.Final_.pdf

¹¹ <https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2022/03/Protocol-for-Reporting-SDG-4.1.1.pdf>

G. UTILISATIONS ET AVANTAGES DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Bien que l'objectif principal de l'établissement de liens entre les politiques pour mesurer les résultats d'apprentissage à l'échelle mondiale soit de relier les évaluations locales, nationales, régionales et internationales aux indicateurs mondiaux, ce processus présente d'autres avantages. Par exemple, comme le montre la **figure 4**, le gouvernement du pays/l'agence d'évaluation et ses partenaires auto-évalueront leur évaluation par rapport aux critères de qualité de ce référentiel. Ces informations pourraient contribuer à améliorer les systèmes éducatifs des pays. Enfin, les résultats de l'atelier de liaison politique devraient aider les pays à identifier le pourcentage et le profil (en supposant que le pays/l'agence d'évaluation ait collecté des informations démographiques sur la population évaluée) des apprenants de leur pays ne répondant pas aux normes de compétence minimales mondiales. Certains pays utilisent ces informations pour mener des études visant à déterminer pourquoi ces lacunes existent et comment ils pourraient au mieux y remédier.

Figure 4 : Processus et avantages de l'établissement de liens entre les politiques



H. UTILISER LE RÉFÉRENTIEL D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Ce référentiel pour l'établissement de liens entre les politiques est conçu pour les équipes du projet, plus particulièrement les animateurs d'ateliers et les personnes ressources (c.-à-d. les responsables gouvernementaux, les agents des agences d'évaluation, les représentants des donateurs et les partenaires) qui organiseront, financeront et/ou mettront en œuvre la méthode dans leur pays ou région.¹² Il contient des lignes directrices pour la mise en œuvre de la méthode.

Le chapitre II décrit le processus d'auto-évaluation pour permettre aux équipes du projet de confirmer que l'évaluation est appropriée pour établir un lien entre les politiques. **Le chapitre III** comprend des

¹² Idéalement, l'évaluation, l'examen ou l'étude du gouvernement utiliserait ce référentiel et cette formation pour mener à bien le processus d'établissement de liens entre les politiques avec ses propres ressources et expertises. Cependant, dans les cas où le gouvernement n'organise pas ce processus de manière autonome, l'organisation responsable et l'équipe de projet doivent travailler en étroite collaboration avec le gouvernement dans la planification et la mise en œuvre du processus afin de garantir l'adhésion et le renforcement des capacités en vue des futurs ateliers.

détails sur la méthodologie d'établissement de liens entre les politiques. **Le chapitre IV** présente des conseils sur la manière de préparer un atelier consacré à l'établissement de liens entre les politiques, notamment sur la manière de sélectionner les animateurs et les participants, à quoi devraient ressembler les invitations, quelle logistique doit être coordonnée, quels supports préparer et comment les préparer, et comment former les animateurs de contenu dans les principales sections de l'atelier. **Le chapitre V** fournit des conseils étape par étape sur la manière de mettre en œuvre un atelier consacré à l'établissement de liens entre les politiques. Enfin, **le chapitre VI** présente les principales considérations pour documenter les résultats de l'atelier consacré à l'établissement de liens entre les politiques et présente des détails sur le processus d'auto-évaluation pour permettre aux équipes du projet de confirmer que les résultats de l'atelier sont suffisamment valables.

La bibliographie contient des références sur l'établissement de liens entre les politiques, l'analyse comparative et d'autres questions psychométriques. Elle comprend le *document de justification de l'établissement de liens entre les politiques (2019)*, qui fournit des informations générales sur la méthode d'établissement de liens entre les politiques, le soutien de la méthode par les donateurs internationaux et des informations sur l'importance de la méthode pour mesurer les résultats en lecture et en mathématiques à l'échelle mondiale.¹³

Les annexes fournissent tous les documents et formulaires nécessaires à l'application des procédures d'établissement de liens entre les politiques décrites dans la boîte à outils. Cela comprend, entre autres, un exemple d'ordre du jour de l'atelier, des modèles de diapositives d'animation, des formulaires d'alignement et d'évaluation des items, un modèle d'évaluation de l'atelier, des formules de calcul des critères de référence et des statistiques, et un plan de rapport technique.

Bien que toutes les personnes impliquées dans l'établissement de liens politiques devraient avoir une compréhension globale du processus de bout en bout, le **tableau 3** contient des guides de navigation qui fournissent des détails sur les sections les plus pertinentes pour chacun des rôles de l'animateur principal, de l'animateur de contenu et d'analyste de données décrits au **chapitre IV, section A**. Il convient de noter que l'animateur principal, l'animateur de contenu et l'analyste de données peuvent faire partie de l'équipe du projet, mais cela dépendra de la manière dont le pays souhaite s'organiser pour mettre en œuvre le projet.

Tableau 3: Guides de navigation

Rôle	Sections principales les plus pertinentes	Annexes les plus pertinentes
Équipe du projet	Tous	Tous
Animateur principal	Chapitre III et Chapitre V	Annexe G, Annexe H, Annexe I, Annexe P, Annexe R, et Annexe S
Animateur de contenu	Chapitre III et Chapitre V	Annexe H, Annexe I, Annexe P, Annexe R, et Annexe S
Analyste de données	Chapitre III	Annexe G, Annexe J, Annexe N, Annexe O, et Annexe T

¹³ Management Systems International (2019). *Policy linking method: Linking assessments to a global standard*. U.S. Agency for International Development (USAID), Washington, D.C.

CHAPITRE II

CHAPITRE II. AUTO-ÉVALUATION DE LA PERTINENCE DE L'ÉVALUATION

A. COLLECTE DES PREUVES ET QUESTIONS À EXAMINER

L'activité d'auto-évaluation sera dirigée par le pays, avec le concours de l'organisation donatrice (le cas échéant) et du partenaire en charge de sa mise en œuvre. Il est essentiel que toutes les preuves, dont l'outil d'évaluation lui-même, les données brutes et la conception de l'échantillonnage, soient partagées avec ceux qui prennent part à l'auto-évaluation.

Pour les évaluations développées à l'aide de la théorie classique des tests (Classical Test Theory, CTT), cela inclura des informations opérationnelles telles que la conception de l'évaluation, la fonctionnalité des items et la discrimination des items. Il comprendra également la répartition des scores pour permettre le développement de données d'impact (voir l'**annexe N**).

Pour les évaluations développées à l'aide de la théorie des réponses aux items (Item Response Theory, IRT), cela inclura des informations opérationnelles telles que le type de modèle IRT utilisé, la conception de l'évaluation, les estimations des paramètres des items (difficulté, discrimination, ajustement, erreur de mesure, etc.), les estimations des capacités des élèves, le plan d'échantillonnage, et toute autre procédure opérationnelle relative au reporting au regard des normes nationales existantes (telles que la correspondance de la probabilité de réponse lors du placement d'un item dans une fourchette de performance). Les informations fournies doivent être suffisantes pour générer les données d'impact requises (voir l'**annexe N**).

Lorsqu'il ne s'agit pas de la langue maternelle des apprenants, les pays doivent tenir compte de la manière dont la langue de l'évaluation peut affecter les résultats de l'établissement de liens entre les politiques. Il est probable que, dans ce cas, les acquis des apprenants soient inférieurs à ce qu'ils auraient été s'ils avaient été évalués dans leur langue maternelle. Cela n'affecte pas la capacité à mener à bien un atelier d'établissement de liens entre les politiques, mais se reflétera dans les rapports au regard de l'ODD 4.1.1

B. CRITÈRES DE VALIDITÉ D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Il existe cinq critères pour cette première auto-évaluation afin de déterminer si l'évaluation est suffisamment valide pour rendre compte de l'ODD 4.1.1. Si une évaluation ne répond pas à ces critères, ils peuvent choisir de continuer avec un atelier consacré à l'établissement de liens entre les politiques, mais les résultats ne seront pas acceptés par l'ISU pour rédiger un rapport mondial.

Pour chaque critère, il existe des exigences minimales essentielles pour qu'une évaluation soit auto-évaluée comme étant suffisamment valable pour le rapport sur les ODD. Dans certains cas, il existe également des exigences souhaitables, notamment des descriptions de ce à quoi ressembleraient des évaluations « bonnes » et « excellentes » pour soutenir le renforcement des capacités. Bien que ces exigences soient souhaitées, les exigences minimales sont les seules essentielles aux rapports sur les ODD.

Les cinq critères portent sur les items suivants :

- **Critère 1** : l'évaluation est-elle suffisamment alignée sur le GPF ?
- **Critère 2** : existe-t-il des preuves que les items de l'évaluation ont été examinés qualitativement et quantitativement pour déterminer leur aptitude à être inclus dans l'évaluation ?
- **Critère 3** : l'échantillon d'apprenants qui a passé l'évaluation est-il représentatif de la population par rapport à laquelle les résultats seront rapportés ?

- **Critère 4** : existe-t-il des preuves que l'évaluation a été administrée de manière standardisée ?
- **Critère 5** : les résultats de l'évaluation sont-ils suffisamment fiables ?

L'équipe de projet doit enregistrer les résultats de son auto-évaluation à l'aide du formulaire de l'**annexe C**.

Critère I – Alignement

Pour s'auto-évaluer par rapport à ce critère, l'équipe du projet devra avoir accès aux items suivants :

- L'outil d'évaluation
- Le cadre ou le cahier des charges de l'évaluation
- Le cadre du programme d'études.

Il existe trois catégories d'alignement, bien qu'elles soient toutes suffisantes pour établir un rapport au regard de l'ODD 4.1.1 :

- **Alignement minimal** – Le contenu de l'évaluation s'aligne sur le nombre minimum de compétences en lecture/mathématiques contenues dans le GPF pour pouvoir faire l'objet d'un rapport au regard de l'ODD 4.1.1, bien que le rapport soit qualifié par une note indiquant le niveau d'alignement.
- **Alignement supplémentaire** – Le contenu de l'évaluation s'aligne sur un nombre minimum de compétences supérieur en lecture/mathématiques dans le GPF pour être adapté à l'établissement de rapports au regard de l'ODD 4.1.1, mais ne répond pas aux exigences d'un fort alignement et sera qualifié comme tel.
- **Fort alignement** – Le contenu de l'évaluation s'aligne fortement sur les compétences en lecture/mathématiques du GPF et convient donc à un rapport sans réserve au regard de l'ODD 4.1.1.

EXIGENCES MINIMALES

Pour établir un rapport au regard de l'ODD 4.1.1, l'évaluation doit être suffisamment alignée sur le GPF. En plus d'une note totale minimale pour l'évaluation, les pays doivent déterminer si l'évaluation est suffisamment approfondie (nombre d'items qui correspondent au moins partiellement à au moins un énoncé de connaissances et/ou de compétences du GPF) et d'étendue (couverture des domaines, concepts et sous-concepts du GPF par au moins un item avec un ajustement partiel) d'alignement. Ceci est déterminé à l'aide d'un exercice d'alignement. Cet exercice est répété par les panélistes lors de l'atelier consacré à l'établissement de liens entre les politiques, mais c'est l'exercice mené lors de l'auto-évaluation qui prime pour déterminer l'adéquation de l'évaluation au rapport. Des exemples supplémentaires pour soutenir le processus d'alignement sont fournis au **chapitre III**.

L'équipe du projet, y compris les experts en contenu appropriés, utilisera la méthode d'alignement Frisbie décrite ici pour compléter les trois sous-étapes suivantes en utilisant le même formulaire d'évaluation de l'alignement qui sera utilisé par les panélistes pendant l'atelier, qui se trouve à l'**annexe H**.

1. Pour chaque item d'évaluation, identifiez les connaissances et/ou compétences dont les apprenants ont besoin pour y répondre correctement.
2. Effectuez une recherche dans le GPF (à l'aide du tableau 3 du GPF) afin de trouver le domaine, le concept, le sous-concept et le ou les énoncés de connaissances et/ou de compétences qui correspondent aux connaissances et/ou aux compétences nécessaires pour répondre

correctement à l’item (pour les évaluations en lecture, examinez également le niveau scolaire du texte, en utilisant les critères d’évaluation de la complexité du texte aux annexes A et B du GPF de lecture).

3. Utilisez l’échelle d’alignement qui suit pour évaluer le niveau d’alignement de l’item.

Échelle d’alignement :

- **Correspondance complète (C)** signifie que tout le contenu requis pour répondre correctement à l’item est contenu dans l’énoncé des connaissances et/ou des compétences, c’est-à-dire que si l’apprenant répond correctement à l’item, c’est parce qu’il utilise complètement les connaissances et/ou les compétences et/ou les compétences décrites dans l’énoncé.
- **Correspondance partielle (P)** signifie qu’une partie du contenu requis pour répondre correctement à l’item est contenue dans l’énoncé des connaissances et/ou des compétences, c’est-à-dire que si l’apprenant répond correctement à l’item, c’est parce qu’il utilise partiellement les connaissances et/ou les compétences décrites dans l’énoncé.
- **Sans correspondance (N)** signifie qu’aucun contenu requis pour répondre correctement à l’item n’est contenu dans les énoncés de connaissances et/ou de compétences, c’est-à-dire que si l’apprenant répond correctement à l’item, c’est parce qu’il n’utilise pas les connaissances et/ou compétences décrites dans le GPF.

Une fois que tous les items ont été notés, l’alignement global sur le GPF est déterminé. Les critères pour les mathématiques sont présentés dans le **tableau 4** et ceux pour la lecture sont présentés dans le **tableau 5**. Lors de la synthèse des résultats au niveau du sous-concept, l’équipe du projet ne doit prendre en compte que les sous-concepts dotés de connaissances et/ou de compétences attendues au niveau scolaire pour lequel l’alignement est effectué (2^{ème} année pour l’ODD 4.1.1(a), 5^{ème} année pour l’ODD 4.1.1(b) et 8^{ème} année pour l’ODD 4.1.1(c)), c.-à-d. ceux qui ont un « x » répertorié dans la colonne de niveau scolaire approprié dans le tableau 3 du GPF.

Tableau 4 : Critères d'alignement des évaluations en mathématiques pour les années 1 à 9

Niveau d'alignement	Catégorie	ODD 4.1.1 (a) GPF 2 ^{ème} année	ODD 4.1.1 (b) GPF niveau 5	ODD 4.1.1 (c) GPF niveau 8
Alignement minimal	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	N (minimum 10 points)		
	Sous-concepts (étendue) :	Points couvrant au moins 2 des 4 sous-concepts N	Points couvrant au moins 5 des 10 sous-concepts N	Points couvrant au moins 4 des 8 sous-concepts N
Alignement supplémentaire	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	N (minimum 10 points) et M et G (minimum 5 points)		
	Sous-concepts (étendue) :	Points couvrant au moins 6 des 11 sous-concepts N, M, et G	Points couvrant au moins 9 des 17 sous-concepts N, M, et G	Points couvrant au moins 7 des 14 sous-concepts N, M, et G
Fortement aligné	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	N (minimum 10 points) et M et G (minimum 5 points) et S et A (minimum 2 points)	N (minimum 10 points) et M et G (minimum 5 points) et S et A (minimum 5 points)	
	Sous-concepts (étendue) :	Points couvrant au moins 7 des 14 sous-concepts	Points couvrant au moins 12 des 21 sous-concepts	Points couvrant au moins 12 des 21 sous-concepts

Légende : N – Nombre et opérations
M – Mesure
G – Géométrie

S – Statistiques et probabilité
A – Algèbre

Tableau 5 : Critères d'alignement des évaluations en lecture pour les années 1 à 9

Niveau d'alignement	Catégorie	ODD 4.1.1 (a) GPF 2 ^{ème} année	ODD 4.1.1 (b) GPF niveau 5	ODD 4.1.1 (c) GPF niveau 8
Alignement minimal	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	D (minimum 10 points) C (minimum 5 points)	R (minimum 10 points)	R (minimum 20 points)
	Sous-concepts (étendue) :	Points couvrant au moins 4 des 7 sous-concepts D&C	Points couvrant au moins 4 des 8 sous-concepts R	Points couvrant au moins 5 des 10 sous-concepts R
Alignement supplémentaire	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	D (minimum 10 points) R (minimum 5 points)	N/D	B1 (minimum 5 points) B2 (minimum 5 points)
	Sous-concepts (étendue) :	Points couvrant au moins 3 des 6 sous-concepts D&R	N/D	Points couvrant au moins 5 des 10 sous-concepts R
Fortement aligné	Durée du test	Score total minimum de 20 si le réglage correspond uniquement au niveau « atteint » Score total minimum de 45 si le réglage correspond aux niveaux « partiellement atteint », « atteint » et « dépasse »		
	Domaine (profondeur) :	R (minimum dix points)	B1 (minimum 5 points) B2 (minimum 5 points)	B1 (minimum 5 points) B2 (minimum 5 points) B3 (minimum 5 points)
	Sous-concepts (étendue) :	Points couvrant au moins 1 des 2 sous-concepts R	Points couvrant au moins 4 des 8 sous-concepts R	Points couvrant au moins 5 des 10 sous-concepts R

Légende : D – Décryptage
C – Compréhension du langage parlé ou signé
R – Compréhension écrite

B1 – Recueillir des informations
B2 – Interpréter des informations
B3 – Réfléchir aux informations

EXIGENCES SOUHAITABLES

Pour toute évaluation de qualité, il est essentiel qu'il y ait un lien clair entre le contenu du programme, ce qui est réellement enseigné et ce qui est évalué. Si les évaluations ne correspondent pas au programme et à ce qui est enseigné en classe, il est peu probable que les élèves obtiennent d'aussi bons résultats lors de l'évaluation.

Bonne note

Pour s'auto-évaluer comme « bon », l'évaluation doit être au moins d'un alignement supplémentaire (les évaluations qui sont alignées de manière minimale ne peuvent pas recevoir la mention « bonne ») et il doit y avoir des preuves claires que l'outil d'évaluation répond aux exigences quantitatives de la spécification ou du modèle de test dans le cadre d'évaluation.

Excellente note

Pour s'auto-évaluer comme « excellent », l'évaluation doit être fortement alignée et il doit y avoir des preuves claires que les cadres d'évaluation et de programme d'études sont alignés.

Les pays doivent disposer d'un cadre du programme d'études¹⁴ qui inclut des détails sur les domaines, les concepts, les sous-concepts et les compétences¹⁵ qui devraient être enseignés en classe par niveau. Les descripteurs doivent être suffisamment détaillés pour indiquer clairement ce qui doit être enseigné. Une activité d'alignement similaire à celle menée avec le GPF peut être menée entre le cadre curriculaire du pays et l'instrument d'évaluation pour déterminer si ceux-ci sont suffisamment alignés.

Critère 2 – Examen de l'item

Pour s'auto-évaluer par rapport à ce critère, l'équipe du projet devra avoir accès aux items suivants :

- Outils d'évaluation, dont des conseils de notation
- Preuves issues du développement de l'outil d'évaluation, qui peuvent prendre la forme d'un rapport technique, de résultats d'analyse de données ou d'entretiens avec les responsables du développement de l'outil d'évaluation
- Statistiques des items.

EXIGENCES MINIMALES

Pour établir un rapport au regard de l'ODD 4.1.1, il faut être en mesure de prouver que les items de l'évaluation ont été examinés quantitativement et qualitativement afin de déterminer s'ils conviennent à l'évaluation.

L'évaluation qualitative doit déterminer si :

- chaque item d'évaluation est considéré comme approprié par les experts concernés pour être inclus dans l'évaluation
- les guides de notation sont cohérents avec ce que l'item est censé mesurer.

L'évaluation qualitative doit déterminer si :

- la difficulté de l'item (par exemple, l'installation de l'item (CTT) ou de l'item sur l'échelle (IRT) est appropriée pour le niveau scolaire

¹⁴ Lorsqu'un pays dispose d'un système hautement décentralisé/fédéralisé en matière de dispositions curriculaires, il doit considérer un petit nombre d'exemples de programmes locaux considérés comme largement représentatifs.

¹⁵ Nous n'attendons pas de tous les pays qu'ils utilisent ces termes dans leurs programmes d'études, mais plutôt qu'ils s'efforcent de détailler les sujets qui doivent être enseignés.

- la discrimination des items (p. ex., l'indice de discrimination pour chaque item est généralement supérieur à 0,2, toutes les exceptions étant rationalisées ou les distracteurs dans des questions à choix multiples doivent être corrélés négativement avec la capacité).

EXIGENCES SOUHAITABLES

Pour toute évaluation de haute qualité, il est important d'examiner les items envisagés pour les inclure dans l'évaluation afin de garantir qu'ils fonctionnent de manière appropriée, y compris pour des sous-groupes de la population (par exemple, les apprenants de genres différents, ceux ayant des besoins éducatifs particuliers et des handicaps, BEPH), ceux de différentes origines ethniques ou culturelles, ceux touchés par des crises et des conflits, ceux des zones rurales et urbaines et ceux qui vivent dans la pauvreté.)

Bonne note

Il n'y a aucune attente supplémentaire concernant ce critère pour justifier une bonne note.

Excellente note

Pour s'auto-évaluer comme « excellent », il faut être en mesure de prouver que les items ont été examinés pour garantir leur pertinence pour les sous-groupes concernés de la population. Ces items probants seront qualitatifs et quantitatifs.

L'évaluation doit paraître exempte de biais. Les items ne doivent pas poser de questions sur des concepts étrangers ou des concepts familiers à certains groupes culturels, ethniques, valides, socio-économiques, de sexe ou géographiques. Par exemple, lire des passages de compréhension qui traitent de fêtes qui ne peuvent être célébrées que par certains groupes ou qui traitent de la neige dans des pays où il ne neige que dans certaines parties du pays serait inapproprié. Des items doivent prouver que les experts ont cherché à savoir si les documents comportaient de tels biais et les ont supprimés le cas échéant.

Les pays doivent également démontrer quelles adaptations de test, le cas échéant, ils ont faites pour les élèves présentant des BEPH.

Les pays doivent également utiliser des techniques statistiques évaluées pour étudier le fonctionnement différentiel des items pour différents sous-groupes de la population, en prenant les mesures nécessaires.

Critère 3 – Échantillon

Pour s'auto-évaluer par rapport à ce critère, l'équipe du projet devra avoir accès aux items suivants :

- Description qualitative de la population pour laquelle ils souhaitent établir un rapport au regard de l'ODD 4.1.1
- Description qualitative de la cohorte à laquelle l'évaluation a été administrée (si différente)
- Informations sur la méthodologie d'échantillonnage. Par exemple, s'il s'agit d'un échantillon aléatoire stratifié, les pays doivent comprendre les détails des strates (qui doivent au moins inclure le district ou d'autres grandes unités administratives).

EXIGENCES MINIMALES

Pour établir un rapport au regard de l'ODD 4.1.1, il faut être en mesure de prouver que l'échantillon d'apprenants qui a passé l'évaluation est représentatif de la population par rapport à laquelle les résultats seront rapportés.¹⁶

¹⁶ Il est admis que pour certains pays, définir ce que signifie « représentatif au niveau national » peut s'avérer difficile en raison du manque de base d'échantillonnage précis. Dans de tels cas, les gouvernements doivent réfléchir à la manière dont ils ont tenté d'obtenir un échantillon approprié et identifier les limites connues de leur approche.

Lorsque l'évaluation est administrée à l'ensemble de la cohorte, l'équipe du projet doit déterminer s'il existe des sous-groupes de la population qui ont été systématiquement exclus. Par exemple, les apprenants non scolarisés, les apprenants vivant dans des zones touchées par des conflits, les apprenants ayant des besoins éducatifs particuliers. Toute exclusion systématique doit être notée pour le rapport.

Lorsque l'évaluation est administrée à un échantillon de la population, la marge d'erreur doit être de 5 % ou moins au niveau de confiance de 95 %.

EXIGENCES SOUHAITABLES

Pour toute évaluation de haute qualité où une approche d'échantillonnage est utilisée, la taille de l'échantillon doit être déterminée de manière à être suffisamment puissante pour permettre aux pays de saisir les changements dans les résultats au fil du temps avec une confiance appropriée.

Bonne note

Il n'y a aucune attente supplémentaire concernant ce critère pour justifier une bonne note.

Excellente note

Pour s'auto-évaluer comme « excellent », la taille minimale de l'effet détectable doit avoir été calculée et réfléchie avant de finaliser les calculs de la taille de l'échantillon afin de garantir que les différences dans le temps puissent être détectées.

Critère 4 – Administration

Pour s'auto-évaluer par rapport à ce critère, l'équipe du projet devra avoir accès aux items suivants :

- Documents d'orientation pour l'administration
- Tous les rapports sur la mise en œuvre des dispositions administratives.

EXIGENCES MINIMALES

Pour établir un rapport au regard de l'ODD 4.1.1, il doit exister des preuves que l'évaluation a été administrée de manière appropriée et standardisée (p. ex., les conditions d'administration sont cohérentes ou la durée d'administration de l'évaluation est respectée). Les guides d'administration doivent être révisés pour plus de clarté et tout incident d'administration inappropriée doit être enregistré. Lorsque des incidents significatifs d'administration inappropriée sont enregistrés, les résultats pertinents doivent être exclus des résultats. Cela nécessitera des vérifications supplémentaires pour confirmer que cela n'affecte pas la représentativité de l'échantillon.

EXIGENCES SOUHAITABLES

Il n'existe aucune exigence supplémentaire souhaitée pour ce critère.

Critère 5 – Fiabilité

Pour s'auto-évaluer par rapport à ce critère, l'équipe du projet devra avoir accès aux items suivants :

- Données de l'administration la plus récente de l'évaluation
- Statistiques de fiabilité calculées à partir de l'analyse des données
- Détails des dispositions relatives à l'assurance qualité pour tous les items notés par l'homme.

EXIGENCES MINIMALES

Pour établir un rapport au regard de l'ODD 4.1.1, la valeur du coefficient alpha¹⁷ (ou statistique de fiabilité équivalente) pour l'évaluation doit être supérieure ou égale à 0,7. En outre, il doit exister des preuves de dispositions appropriées en matière d'assurance qualité pour tout item évalué par des humains. Cela pourrait inclure la notation des items avec un score convenu à l'avance ou une double notation d'un échantillon de réponses.

EXIGENCES SOUHAITABLES

Pour toute évaluation de haute qualité, il est essentiel que les résultats de l'évaluation présentent des niveaux de fiabilité appropriés, ce qui signifie que si le test était à nouveau administré à un autre échantillon d'étudiants avec un ensemble différent d'enquêteurs ou de surveillants de test, les résultats seraient similaires. Les résultats d'évaluation qui fluctuent considérablement en fonction de la personne qui administre l'évaluation, de la personne qui l'a notée ou des étudiants spécifiques d'un échantillon qui passent l'évaluation ne sont pas des évaluations fiables.

Bonne note

Pour s'auto-évaluer comme « bon », le niveau d'accord entre les évaluateurs humains et les scores préalablement convenus, ou les scores doublement notés, doit être supérieur à 80 %.

Excellente note

Pour s'auto-évaluer comme « excellent », le pays doit communiquer d'autres mesures de fiabilité de l'évaluation, par exemple la constance de la classification, l'exactitude de la classification ou la fiabilité entre les évaluateurs, avec des niveaux conformes aux meilleures pratiques internationales.

Note d'auto-évaluation globale

Pour pouvoir rendre compte des résultats au regard de l'ODD 4.1.1, les pays doivent s'auto-évaluer selon les exigences minimales pour chacun des critères.

Pour s'auto-évaluer comme « bon », les pays doivent en outre s'auto-évaluer au niveau « bon » pour les critères 1 et 5. Pour s'auto-évaluer comme « excellent », les pays doivent s'auto-évaluer au niveau « excellent » pour les cinq critères.

C. ÉTAPES SUIVANTES

Si les pays s'auto-évaluent comme satisfaisant aux exigences minimales (ou comme « bons » ou « excellents »), ils peuvent continuer à organiser un atelier sur l'établissement de liens entre les politiques. Les résultats de l'atelier pourront faire l'objet d'un rapport au regard de l'ODD 4.1.1 tant que l'atelier est mené conformément aux orientations dudit référentiel. Cela sera auto-évalué à la fin du processus, comme décrit au **chapitre VI**.

Les pays qui s'auto-évaluent comme ne répondant pas aux exigences minimales peuvent choisir de poursuivre un atelier consacré à l'établissement de liens entre les politiques à des fins de renforcement des capacités, mais les résultats ne seront pas acceptés par l'ISU pour établir un rapport au regard des objectifs de l'ODD 4.1.1. Ils souhaiteront peut-être également examiner s'il est possible de modifier leurs modalités d'évaluation actuelles pour les rendre compatibles avec la méthode d'établissement de liens entre les politiques. La manière d'y parvenir dépendra des critères auxquels ils ne répondront pas.

¹⁷ Également connu sous le nom de coefficient alpha de Cronbach

S'ils n'ont pas satisfait au critère 1 sur l'alignement sur le GPF, ils voudront peut-être envisager les changements qui seraient nécessaires à leur cadre d'évaluation. Par exemple, inclure une plus grande variété d'items numériques des différents sous-concepts ou inclure davantage d'items de compréhension en lecture, pour s'assurer que leur évaluation est alignée.

S'ils n'ont pas satisfait au critère 2 sur l'examen des items, ils voudront peut-être revoir leurs modalités d'élaboration des tests afin de s'assurer de leur conformité à la pratique internationale. Par exemple, ils pourraient mettre en œuvre des processus tels qu'énoncés dans les Normes pour les tests éducatifs et psychologiques (2014), qui incluraient l'examen des items avant de les intégrer à une évaluation.

S'ils ne satisfont pas au critère 3 sur l'échantillonnage, ils voudront peut-être revoir leur méthodologie d'échantillonnage pour s'assurer qu'elle produit un échantillon représentatif au niveau national, par exemple en excluant moins d'apprenants ou en garantissant une couverture régionale dans l'échantillon.

S'ils ne satisfont pas au critère 4 sur l'administration, ils peuvent souhaiter revoir leurs directives d'administration ou proposer aux administrateurs de test davantage de formations pour assurer la cohérence.

S'ils ne satisfont pas au critère 5, ils souhaiteront peut-être améliorer leurs dispositions d'assurance qualité pour les items notés par des humains ou travailler avec un partenaire technique pour déterminer les moyens d'améliorer la fiabilité de leur évaluation.

CHAPITRE III

CHAPITRE III : LA MÉTHODE D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Une fois que l'évaluation a été confirmée comme étant appropriée pour l'établissement de liens entre les politiques, l'équipe du projet doit confirmer quels items seront utilisés dans l'atelier. Pour les évaluations où le même outil est administré à tous les élèves, l'équipe du projet peut décider d'utiliser tous les items. Toutefois, si l'évaluation est très longue, comporte plusieurs formulaires ou un plan d'échantillonnage complexe, l'équipe du projet devra déterminer quels items elle utilisera en fonction des jours disponibles pour l'atelier et d'autres critères.

La méthode d'établissement de liens entre les politiques commence par un examen approfondi des principaux documents qui constituent la base de l'atelier – le GPF et la ou les évaluations étant liées au GPF et à l'ODD 4.1.1. Suite à cette familiarisation, les animateurs guident les panélistes à travers trois tâches principales :

- **Tâche 1** – vérifier l'alignement du contenu entre la ou les évaluations et le GPF à l'aide d'une procédure standardisée
- **Tâche 2** – faire correspondre les items d'évaluation avec le GPF, c'est-à-dire les GPL et les GPD
- **Tâche 3** – définir trois critères de référence mondiaux pour chaque évaluation en utilisant une méthode standardisée (une version modifiée de la méthodologie Angoff) à travers deux séries de notations

Chacune de ces activités est décrite en détail plus loin dans ce chapitre.

A. CHOIX DES ITEMS

Pour rendre l'atelier consacré à l'établissement de liens entre les politiques gérable, il est recommandé que pour les évaluations non chronométrées, un maximum de 45 points soient utilisés dans le processus (ce qui signifie 45 items si chacun vaut un point de score, mais moins s'il y a des items polytomiques). Comme indiqué dans la section d'auto-évaluation, cela suffit pour définir les trois GPL (« atteint », « partiellement atteint » et « dépasse ») pour un seul niveau scolaire. Moins d'items peuvent être sélectionnés si seul le GPL « atteint » est défini via l'établissement de liens entre les politiques, bien qu'un minimum de 20 points soit nécessaire. Pour les évaluations chronométrées, il sera possible d'inclure un plus grand nombre de points en fonction de la nature de la tâche (par exemple, le nombre de points peut être lié au nombre de mots lus dans un délai imparti). L'équipe du projet doit déterminer si l'atelier durera suffisamment longtemps pour que les panélistes puissent prendre tous les jugements nécessaires lors de la détermination du nombre d'items à utiliser dans une évaluation chronométrée.

Si l'évaluation dépasse 45 points (pour une évaluation non chronométrée) ou si un plan d'échantillonnage complexe est utilisé, l'équipe du projet devra alors sélectionner les items à utiliser. Si l'évaluation est administrée dans plusieurs langues, les items sélectionnés doivent être dans la langue dans laquelle l'atelier sera dispensé. Les liens avec d'autres langues doivent être établis de manière statistique, lorsque cela est possible, ou en organisant des ateliers distincts dans chaque langue d'administration.

Lors de la sélection des items, l'équipe du projet doit considérer les items suivants :

- **Couverture du contenu** – les items sélectionnés doivent refléter largement la couverture du domaine, des concepts et des sous-concepts de l'ensemble du test ou de l'ensemble des items. Par exemple, si 60 % de l'ensemble des tests/items correspondent au concept « recueillir des informations » dans le domaine de compréhension écrite, alors 60 % des items sélectionnés pour l'établissement de liens entre les politiques doivent également être alignés sur l'item « recueillir des informations ».

- **Alignement GPL** – les items sélectionnés doivent refléter globalement les différents GPL définis lors de l’atelier consacré à l’établissement de liens entre les politiques. Par exemple, si les trois GPL sont définis au cours de l’atelier, alors il devrait y avoir 15 items alignés sur chacun des GPL, tels qu’identifiés lors de la tâche d’alignement d’auto-évaluation. Si seul le niveau « atteint » au GPL est défini, alors les 20 items doivent principalement être alignés sur le niveau « atteint » au GPL dans le GPF, bien que certains items « partiellement » et « dépasse » puissent également être inclus.
- **Fonctionnement des items** – les items sélectionnés doivent être ceux les plus fonctionnels dans l’évaluation. Par exemple, s’il existe des items présentant une discrimination négative ou présentant des items différentiels fonctionnant en faveur d’un sexe particulier, ceux-ci doivent être supprimés.
- **Items classés comme « sans correspondance » lors de l’exercice d’alignement** – les items sélectionnés devraient idéalement tous correspondre au moins partiellement au GPF. Étant donné que l’exercice d’alignement d’auto-évaluation a déjà déterminé que l’évaluation est suffisamment alignée sur le GPF pour soutenir l’établissement de liens entre les politiques, il est utile de supprimer les items non alignés pour éviter d’avoir à les traiter lors de l’atelier. Par exemple, lorsque cela est possible, les items évaluant la grammaire, la ponctuation et l’orthographe doivent être supprimés de l’ensemble des tests/items de l’atelier. Le niveau d’alignement sur le GPF pour le reporting utilisera le résultat de l’exercice d’alignement d’auto-évaluation qui inclut les items classés comme « sans correspondance ».
- **Difficulté des items (applicable uniquement lorsque l’IRT est utilisé)** – lorsqu’une évaluation intègre une rotation des items sur différentes formes de test (une approche d’échantillonnage matriciel des items), l’ensemble des items sélectionnés doit être aussi espacé que possible, du plus facile au plus difficile sur l’échelle IRT sous-jacente. Les items les plus appropriés à utiliser pour établir des liens entre les politiques pourraient ne pas tous être contenus dans un seul formulaire de test existant. En revanche, l’ensemble d’items sélectionnés utilisera tous les items d’évaluation pour fournir la meilleure discrimination et couverture possible des connaissances et des compétences par les items inclus dans l’atelier de liaison politique.

B. FAMILIARISATION

Pour entreprendre avec succès l’atelier consacré à l’établissement de liens entre les politiques, il est essentiel que les panélistes connaissent le GPF et l’outil d’évaluation. Cette familiarisation peut avoir lieu avant ou au début de l’atelier en fonction du temps disponible.

Pour le GPF, la familiarisation doit commencer par expliquer la terminologie nécessaire (GPL, GPD, domaine, concept, sous-concept) et la manière dont le GPF est structuré. Les panélistes doivent ensuite examiner le GPF pour l’année pertinente (2^{ème} année pour l’ODD 4.1.1(a), 5e année pour l’ODD 4.1.1(b) et 8e année pour l’ODD 4.1.1(c) afin de leur permettre d’intérioriser les normes. Dans la mesure du possible, les panélistes devraient s’engager dans des tâches actives pour étayer leur compréhension du GPF plutôt que de simplement le lire.

Pour l’outil d’évaluation, la familiarisation doit expliquer comment l’évaluation est structurée et administrée. Dans la mesure du possible, les panélistes doivent avoir la possibilité d’administrer l’évaluation, ou du moins de la passer eux-mêmes.

C. TÂCHE 1 - ALIGNER L’ÉVALUATION SUR LE GPF

L’alignement est le processus consistant à vérifier si un certain item d’une évaluation utilisée pour l’établissement de liens entre les politiques correspond à un sous-concept/énoncé de connaissances et/ou de compétences dans le GPF. Il est important de distinguer l’activité d’alignement de la tâche 1 du travail d’alignement mené par le gouvernement/l’agence d’évaluation dans le cadre de l’auto-évaluation. L’exercice d’alignement préalable à l’atelier vise à garantir qu’il y a un alignement suffisant entre

l'évaluation du pays/de l'agence d'évaluation et le GPF pour procéder à l'établissement de liens entre les politiques. En revanche, au cours de l'atelier, l'activité d'alignement vise à familiariser davantage les panélistes avec le GPF, en particulier les connaissances et compétences couvertes par celui-ci, et à générer des notes des panélistes sur la profondeur et l'étendue de l'alignement entre les évaluations et le GPF. Cette compréhension aidera les panélistes dans le processus d'analyse comparative qui se déroule dans la tâche 3, car il s'agit de la première étape pour préciser les attentes du GPF que mesurent les évaluations. La tâche 1 comporte deux étapes :

1. Les panélistes notent la concordance entre l'évaluation liée et le GPF
2. Les animateurs de l'atelier et l'analyste de données synthétisent les résultats de l'activité d'alignement (les rôles et les responsabilités sont décrits plus en détail ci-dessous)

Étape 1 - Exercice d'alignement des panélistes

À l'étape 1, après avoir reçu des instructions sur la tâche et avoir travaillé sur quelques exemples avec les animateurs, les panélistes doivent travailler de manière indépendante, en procédant item par item en utilisant la méthode d'alignement Frisbie décrite ici pour compléter les trois sous-étapes suivantes à l'aide du formulaire d'évaluation de l'alignement, qui se trouve à **l'annexe H**.

1. Pour chaque item d'évaluation, identifiez les connaissances et/ou compétences dont les apprenants ont besoin pour y répondre correctement
2. Effectuez une recherche dans le GPF (à l'aide du tableau 3 du GPF) pour trouver le domaine, le concept, le sous-concept et les énoncés de connaissances et/ou de compétences qui correspondent aux connaissances et/ou compétences nécessaires pour répondre correctement à l'item (pour les évaluations en lecture, examinez également le niveau scolaire du texte, en utilisant les critères d'évaluation de la complexité du texte dans les annexes A et B du GPF de lecture)
3. Utilisez l'échelle d'alignement qui suit pour évaluer le niveau d'alignement de l'item

ÉCHELLE D'ALIGNEMENT :

- **Correspondance complète (C)** signifie que **tout le contenu** requis pour répondre correctement à l'item est contenu dans l'énoncé des connaissances et/ou des compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il utilise complètement les connaissances et/ou les compétences et/ou les compétences décrites dans l'énoncé.
- **Correspondance partielle (P)** signifie qu'une **partie du contenu** requis pour répondre correctement à l'item est contenue dans l'énoncé des connaissances et/ou des compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il utilise partiellement les connaissances et/ou les compétences décrites dans l'énoncé.
- **Sans correspondance (N)** signifie **qu'aucun contenu** requis pour répondre correctement à l'item n'est contenu dans les énoncés de connaissances et/ou de compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il n'utilise pas les connaissances et/ou les compétences décrites dans le GPF.

De plus amples détails sur l'échelle apparaissent dans la **Figure 5** ci-dessous.

Figure 5 : Échelle d'alignement et nombre d'énoncés de connaissances et/ou de compétences auxquels un item s'aligne

Si un item a une évaluation de **correspondance complète (C)** avec un énoncé spécifique de connaissances et/ou de compétences, les panélistes ne doivent pas l'associer à d'autres énoncés de connaissance et/ou de compétence, ce qui signifie qu'il est aligné uniquement avec un énoncé dans le GPF.

Si un item a une évaluation de **correspondance partielle (P)** avec un énoncé spécifique de connaissances et/ou de compétences, les panélistes doivent généralement l'associer à un ou deux énoncés supplémentaires de connaissances et/ou de compétences contenus dans le GPF.

Si un item reçoit la mention **sans correspondance (N)** avec des énoncés de connaissances et/ou de compétences, les panélistes ne doivent pas le faire correspondre à des énoncés de connaissances et/ou de compétences.

Un exemple d’item avec « correspondance complète » est donné dans la **figure 6** avec un item qui demande à un apprenant comment huit cent soixante-dix est écrit sous forme standard. Dans cet exemple, le panéliste a identifié que la connaissance ou la compétence nécessaire pour répondre correctement à cet item est la capacité de lire et d’écrire des nombres entiers jusqu’à 1 000. Cette compétence est couverte dans le GPF sous le domaine « connaissance des nombres », le concept « nombre entier » et le sous-concept « identifier et compter en nombres entiers ». Enfin, le panéliste a évalué cet alignement comme « correspondance complète », puisque toutes les connaissances et/ou compétences nécessaires pour répondre correctement à cet item sont contenues dans cet énoncé unique de connaissances et/ou de compétences.

Figure 6 : Exemple d’alignement d’un item avec le GPF et avec correspondance complète

1. Comment huit cent soixante-dix est-il écrit sous une forme standard ?

- A. 807
- B. 870
- C. 817
- D. 871

Domaine : Nombre et opérations

Concept : Nombres entiers

Sous-concept : Identifier, compter et repérer la grandeur relative des nombres entiers

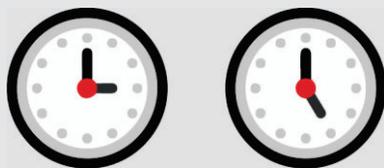
Connaissances ou compétences (norme de contenu) : Compter, lire et écrire des nombres entiers

Correspondance : Pour répondre correctement à cet item, l’apprenant doit être capable d’identifier et de compter des nombres entiers. Par conséquent, l’item peut être évalué comme « correspondance complète » avec l’énoncé de connaissance et/ou de compétence(s) car il ne nécessite que les connaissances ou compétences de cet énoncé unique.

Un exemple d’item avec « correspondance partielle » est présenté dans la **figure 7**. Le panéliste a évalué cet outil comme étant en correspondance partielle, car pour répondre correctement à cet item, un apprenant aurait besoin de connaissances ou de compétences décrites par deux énoncés différents de connaissances et/ou de compétences.

Figure 7 : Exemple d’alignement d’un item sur le GPF avec correspondance partielle

Quelle est la différence d’heure affichée entre ces deux horloges ?



Domaine : Mesure

Concept : Temps

Sous-concept : lire l’heure ET résoudre des problèmes impliquant le temps

Connaissance ou compétence (norme de contenu) : lire l’heure ET résoudre des problèmes impliquant le temps

Correspondance : correspondance partielle car cela nécessite les connaissances et/ou les compétences de deux normes de contenu.

Un exemple d’item « sans correspondance » apparaît dans la **figure 8**. Le panéliste a jugé cet item comme étant sans correspondance, car pour y répondre correctement, un apprenant aurait besoin de connaissances ou de compétences qui ne sont pas décrites dans le GPF. Idéalement, les items sans correspondance avec le GPF auraient été supprimés lors de l’étape de sélection des items. Si cela n’est pas

possible, des instructions seront données sur la manière de traiter les items sans correspondance aux étapes ultérieures du processus.

Figure 8 : Exemple d’alignement d’un item sur le GPF sans correspondance

<p>Laquelle de ces phrases est ponctuée correctement comme une question ?</p> <p>A. Où est le chat !</p> <p>B. Où est le chat.</p> <p>C. Où est le chat ?</p> <p>D. Où est le chat :</p>	<p>Domaine : Sans objet</p> <p>Concept : Sans objet</p> <p>Sous-concept : Sans objet</p> <p>Connaissances ou compétences (norme de contenu) : Sans objet</p>
--	--

Correspondance : sans correspondance puisque les connaissances requises pour répondre à l’item concernent une ponctuation qui n’est pas référencée dans le GPF.

Étape 2 : Résumé des résultats des animateurs

Une fois que tous les panélistes ont terminé leur tâche d’alignement, les animateurs doivent résumer les résultats en faisant la moyenne du nombre d’items que les panélistes ont alignés sur chaque domaine, concept et sous-concept. Même si l’alignement se produit au niveau des connaissances et/ou des compétences, les critères d’alignement se situent au niveau des sous-concepts. En tant que tels, les animateurs doivent résumer les résultats jusqu’au niveau du sous-concept. Les items avec correspondance complète et partielle comptent pour l’alignement, mais chaque item ne doit être compté qu’une seule fois, même s’il s’agit d’une correspondance partielle (remarque : pour les items qui correspondent partiellement, à des fins de synthèse, les animateurs doivent compter le domaine, le concept et le sous-concept qu’ils estiment décrire le mieux les connaissances et/ou compétences les plus importantes nécessaires pour répondre correctement à l’item. Cela pourrait être déterminé en utilisant les résultats de l’exercice d’alignement initial entrepris dans le cadre du processus d’auto-évaluation). Étant donné que le but de l’activité d’alignement est de familiariser davantage les panélistes avec le GPF et l’évaluation, elle n’est pas essentielle pour que les panélistes

à ce stade, soient d’accord les uns avec les autres ou avec l’exercice d’alignement initial. En supposant que l’exercice d’alignement initial ait indiqué un alignement suffisant, le résultat de cet exercice devrait renforcer la confiance dans la pertinence de l’établissement de liens entre les politiques.

Un exemple de résultats sommaires pour une évaluation de 3e année comportant 26 items apparaît ci-dessous. Il convient de noter qu’étant donné que l’établissement de liens entre les politiques pour cette évaluation serait lié à la 2^{ème} année dans le GPF pour les rapports sur l’ODD 4.1.1 (a), les domaines, les concepts et les sous-concepts indiqués ne concernent que ceux inclus dans la 2^{ème} année (comme indiqué dans le tableau 3 du GPF).

Tableau 6 : Exemple de résumé des résultats d’alignement pour une évaluation de 3e année

Domaine		Items
N	Nombre et opérations	14
M	Mesure	7
G	Géométrie	3
S	Statistiques et probabilité	2
A	Algèbre	0
Total		26

Concept		Items
N1	Nombres entiers	14
M1	Longueur, poids, capacité, volume, surface et périmètre	3
M2	Heure	4
M3	Devise	0
G1	Propriétés des formes et des figures	2
G2	Visualisations spatiales	0
G3	Position et orientation	1
S1	Gestion des données	2
A1	Schémas	0
A3	Relations et fonctions	0
Total		26

Sous-concept		Items
N1.1	Identifier et compter en nombres entiers, et repérer leur grandeur relative	4
N1.2	Représenter des nombres entiers de manière équivalente	0
N1.3	Résoudre des opérations en utilisant des nombres entiers	8
N1.4	Résoudre des problèmes réels impliquant des nombres entiers	2
M1.1	Utiliser des unités standard et non standard pour mesurer, comparer et classer	3
M2.1	Donner l'heure	2
M2.2	Résoudre des problèmes impliquant l'heure	2
M3.1	Utiliser différentes unités monétaires pour créer des montants	0
G1.1	Reconnaître et décrire des formes et des figures	2
G2.1	Composer et décomposer des formes et des figures	0
G3.1	Décrire la position et l'orientation d'objets dans l'espace	1
S1.1	Recueillir et interpréter des données présentées à l'écran	2
A1.1	Reconnaître, décrire, développer et générer des modèles	0
A3.2	Démontrer une compréhension des équivalences	0
Total		26

Les animateurs peuvent choisir de partager le niveau d'alignement de l'évaluation avec le GPF en termes de critères utilisés pour l'auto-évaluation (voir **chapitre II**), en se concentrant à la fois sur la profondeur (nombre d'items qui correspondent au moins partiellement à au moins un énoncé de connaissances et/ou de compétences du GPF) et l'étendue (couverture des domaines, des concepts et des sous-concepts du GPF par au moins un item avec une correspondance partielle) de l'alignement. Lorsque l'on note un désaccord important entre les résultats de l'alignement initial et celui des panélistes, les animateurs voudront se concentrer sur ce point lors de la tâche suivante, où le consensus est l'objectif.

D'après les critères mentionnés dans le **chapitre II**, il est clair que l'exemple d'évaluation de 3^e année décrit dans le **tableau 6** serait considéré comme « fortement aligné » puisqu'il : 1) contient au moins dix items numériques (14 au total), au moins cinq au total en matière d'items de mesure et de géométrie (10 au total), et au moins deux items de probabilités et statistiques et algèbre (2 au total) et 2) comportent des items couvrant au moins 7 des 14 sous-concepts avec les connaissances et/ou compétences attendues en 2^{ème} année (9 sous-concepts sur 14 sont couverts).

D. TÂCHE 2 – FAIRE CORRESPONDRE LES ITEMS D'ÉVALUATION AVEC LES GPL ET LES GPD

La tâche 2 s'appuie sur la compréhension des panélistes des items d'évaluation et du GPF acquis grâce à l'activité d'alignement. La mise en correspondance est le processus consistant à trouver le descripteur mondial de compétence (GPD) spécifique au niveau scolaire approprié (2^{ème} année pour l'ODD 4.1.1 (a), 5^{ème} année pour l'ODD 4.1.1 (b) et 8^e année pour l'ODD 4.1.1 (c) dans le GPF qui correspond à un certain item d'une évaluation utilisée pour l'établissement de liens entre les politiques. L'objectif de la tâche 2 est de mieux cerner les attentes des apprenants mesurées par chaque item d'évaluation. Cela

aidera les panélistes à savoir quel GPD (norme de performance) ils devraient prendre en compte lorsqu'ils évaluent si un apprenant peu compétent répondrait correctement à l'item dans le processus d'analyse comparative (tâche 3). Dans cette tâche, les panélistes sont invités à faire passer leur travail d'alignement au niveau supérieur en faisant correspondre chaque item aux GPL et GPD appropriés dans le GPF.¹⁸ Lorsque des divergences d'opinions se présentent lors de l'exercice d'alignement, celles-ci doivent être discutées. Les animateurs doivent se référer à l'exercice d'alignement initial entrepris dans le cadre de l'auto-évaluation pour appuyer la discussion. Les panélistes doivent travailler en groupes pour parvenir à un consensus sur les réponses aux trois questions suivantes pour chaque item d'évaluation :

1. **Quelles connaissances et/ou compétences sont requises pour répondre correctement aux items ?** Les panélistes peuvent s'appuyer sur leur travail à ce sujet au cours de la tâche 1, comparer les réponses et parvenir à un consensus.
2. **Qu'est-ce qui rend l'item simple ou complexe ?** Dans cette étape, les panélistes doivent considérer des items tels que : les distracteurs (parmi les options à choix multiples), si la langue utilisée pour poser la question est la langue que l'apprenant est habitué à entendre en classe, si le sujet (pour un passage de lecture) est susceptible d'être familier, et si les images incluses dans l'item sont susceptibles d'être familières pour l'apprenant et similaires ou différentes de celles présentées dans le matériel pédagogique. Par exemple, dans l'exemple fourni dans la **Figure 9** ci-dessous, le panéliste pourrait dire que le fait que la question utilise exactement les mêmes mots que ceux utilisés dans la première phrase du passage rend facile cet item. Une chose qui pourrait rendre les choses difficiles serait que les apprenants ne soient pas familiers avec les chiens parce qu'ils n'existent pas dans leur contexte.
3. **Quelle est le GPL le plus bas le plus appropriée pour l'item ?** Les panélistes doivent lire les GPD pour chaque GPL au niveau scolaire (et les niveaux inférieurs) pour déterminer quels GPL et GPD correspondent le mieux à un niveau scolaire. Ils doivent sélectionner le GPL le plus bas qui correspond aux connaissances et/ou compétences dont les apprenants ont besoin pour répondre correctement à l'item. Si l'item correspond à plus d'un énoncé de connaissances et/ou de compétences (tel que déterminé dans la tâche 1) et, par conséquent, à plus d'un GPD, le panéliste doit sélectionner le plus élevé des GPL, car un apprenant ne serait pas en mesure de répondre à l'item sans les connaissances et/ou compétences décrites dans ce GPD. Si l'item est trop difficile à faire correspondre au niveau scolaire pour lequel les critères sont fixés, les panélistes doivent noter que l'item dépasse le niveau. Pour les évaluations en lecture il est important de noter que, les panélistes devront souvent évaluer le niveau scolaire du décodage, de la compréhension écrite ou de la compréhension du passage en langue parlée ou signée, car de nombreux GPD sont les mêmes d'une année à l'autre avec la seule différence étant le niveau scolaire du passage. Les annexes A et B du GPF Lecture contiennent des critères et des exemples pour aider les panélistes à évaluer le niveau scolaire du passage.

La **figure 9** fournit un exemple tiré des diapositives de facilitation d'atelier incluses à **l'annexe G**. Dans cet exemple d'item, les apprenants sont invités à lire le passage suivant :

Van est à l'école. Il a de nouveaux crayons.

Van dessine un grand arbre aux feuilles vertes et aux fleurs rouges.

¹⁸ Notez que si au cours de l'étape 1, 2 ou 3, le gouvernement décide qu'il souhaite uniquement fixer un critère de référence pour le niveau satisfaisant ou si le gouvernement/l'agence d'évaluation ou le comité d'examen de l'objectif 4.1.1 décide que l'évaluation est trop courte pour accueillir trois points de repère aux trois principaux GPL, les panélistes n'ont alors qu'à correspondre au GPD du niveau scolaire plutôt qu'au GPL.

Les apprenants sont ensuite invités à répondre à la question « Où est Van ? » Cette question correspond à l'énoncé de connaissances ou de compétences consistant à récupérer une seule information explicite à partir d'un texte de niveau scolaire par correspondance directe de mots. Le panéliste a identifié ce qui rend cet item facile ou difficile dans la case supérieure de cet exemple. Étant donné que le GPF Lecture nécessite une évaluation du niveau scolaire du passage (à l'aide de l'annexe C du GPF), les panélistes doivent déterminer le niveau du passage avant d'identifier le GPL et le GPD. Dans cet exemple, le panéliste a déterminé que le passage est un passage à deux niveaux. En conséquence, l'item s'aligne sur le niveau Partiellement atteint les compétences minimales mondiales en deuxième année.

Figure 9 : Exemple de correspondance d'items avec les GPL et les GPD

Facile ou difficile : une chose qui rend la question facile est qu'elle utilise la même formulation que le passage. Les deux contiennent le mot « est ». De plus, Van est un nom courant dans ce contexte.

Domaine : compréhension écrite

Concept : recueillir des informations

Sous-concept : localiser les informations explicitement indiquées

Niveau de passage : 2^{ème} année

Connaissance ou

compétence : extraire une seule information explicite à partir d'un texte continu au niveau scolaire par correspondance directe ou rapprochée de mots

GPL et GPD (norme de performance) :

Partiellement atteint : extraire un seul item d'information explicite et important à partir d'un texte de 2^{ème} année par correspondance directe ou rapprochée lorsque les informations requises sont adjacentes au mot correspondant et qu'il n'y a aucune information concurrente.

Atteint : extraire une seule information explicite à partir d'un texte de 2^{ème} année par correspondance directe ou rapprochée de mots lorsque les informations requises sont adjacentes au mot correspondant et que les informations concurrentes sont limitées.

Dépasse : extraire une seule information explicite à partir d'un texte continu au niveau scolaire de 2^{ème} année par correspondance directe ou rapprochée de mots.

Une fois ce processus de mise en correspondance terminé, les animateurs demandent aux panélistes de se concentrer sur la mise en correspondance avec les GPD qui correspondent aux items. Les panélistes doivent enregistrer les réponses de leur groupe aux trois questions posées dans cette tâche directement à côté de chaque item de leur livret de test/outil d'évaluation.

Figure 10 : Items correspondants identifiés comme « sans correspondance » avec le GPF

Pour les items qui ont été déterminés comme étant « sans correspondance » avec le GPF, et pour lesquels, il n'a pas été possible de les retirer du processus, la discussion devrait se concentrer sur la tentative de déterminer si l'item est le plus approprié pour les apprenants au niveau partiellement atteint, atteint ou dépasse les normes de performance. Pour ce faire, les panélistes doivent suivre ces étapes :

- Imaginez un groupe d'apprenants qui sont mieux décrits par les GPD dans le niveau « partiellement atteint ».
- En utilisant leur expérience d'enseignement à de tels apprenants, déterminez s'il s'agit d'un item approprié pour ces apprenants et s'ils seraient susceptibles d'y répondre correctement.
- Si l'item est jugé approprié pour les apprenants au niveau partiellement atteint, cela peut être enregistré.
- Dans le cas contraire, le processus doit être répété pour le niveau « atteint », puis pour le niveau « dépasse », si nécessaire.
- Si l'item est jugé trop difficile pour être noté, il doit alors être enregistré comme étant au-dessus du niveau supérieur.

E. TÂCHE 3 : LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DE CRITÈRES DE RÉFÉRENCE

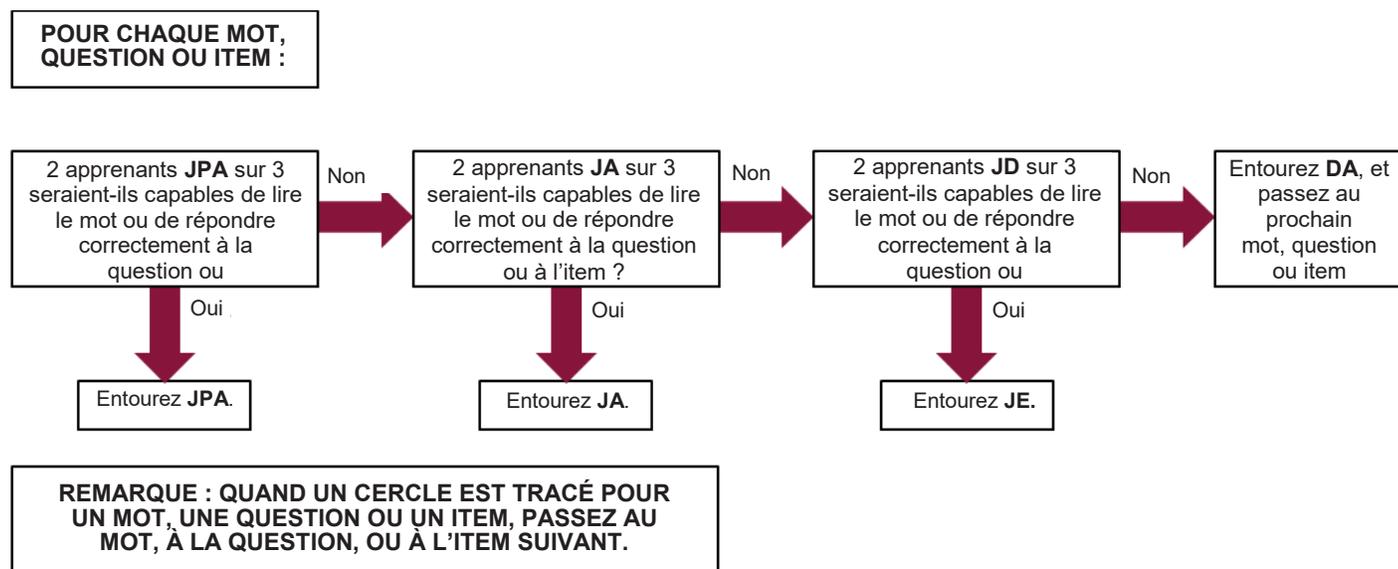
La tâche 3 est la tâche la plus importante de l'atelier consacré à l'établissement de liens entre les politiques, car c'est ici que les panélistes établissent des critères de référence en évaluant la façon dont les apprenants dont les capacités en lecture ou en mathématiques correspondent aux connaissances et/ou compétences alignées sur chaque item de la tâche 1 et comment les GPD correspondant à chaque item de la tâche 2 fonctionneraient sur chaque item. La tâche 3 s'appuie sur la méthode Angoff pour établir des critères. La méthode Angoff est une méthode centrée sur les tests qui convient aux différents types d'évaluations administrées dans différents pays. Avec la méthode Angoff Oui-Non, les panélistes doivent utiliser un formulaire de notation d'items (voir **Annexe I**) pour noter chacun des items des outils d'évaluation, en suivant les quatre étapes suivantes :

- **Étape 1** : identifier ou conceptualiser trois apprenants aux compétences minimales à chaque GPL décrit par le GPF (2^{ème} année pour l'ODD 4.1.1(a), 5^{ème} année pour l'ODD 4.1.1(b) et 8^{ème} année pour l'ODD 4.1.1(c)).¹⁹ Les apprenants peu compétents sont ceux qui obtiennent des résultats égaux ou légèrement supérieurs aux GPD qui décrivent le GPL. Estimer les performances de ces apprenants à chacun des items d'évaluation. Ces apprenants sont appelés apprenants juste partiellement atteint (JPA), juste atteint (JA) et juste dépasse (JD).
- **Étape 2** : procédez item par item en l'examinant et en identifiant les connaissances et/ou compétences requises pour y répondre correctement. L'idée est de se concentrer sur le contenu de l'item en relation avec le ou les énoncés de connaissances et/ou de compétences du GPF. Réfléchissez à ce qui fait la facilité ou la complexité de l'item (p. ex., la formulation de la racine de l'item et la force des options ou des distracteurs incorrects) et quels types d'erreurs peuvent être possibles ou raisonnables (Remarque : les panélistes doivent avoir enregistré ces informations sur leur livret de test/outil d'évaluation au cours de la tâche 2).
- **Étape 3** : sélectionnez le GPL le plus bas, avec le GPD associé, pour les connaissances et/ou compétences nécessaires pour répondre correctement à l'item (les panélistes doivent avoir enregistré ces informations sur leur livret de test/outil d'évaluation au cours de la tâche 2). Lorsque l'item est classé comme « sans correspondance » (et qu'ils n'ont pas été supprimés de l'ensemble des items utilisés), les panélistes doivent utiliser la discussion de l'exercice de correspondance pour déterminer le GPL le plus bas qui répondrait correctement à l'item, même s'il n'est pas lié au GPF.
- **Étape 4** : sur la base d'une compréhension des étapes 1 à 3, suivez la procédure présentée dans l'organigramme de la **Figure II** ci-dessous, qui permet aux panélistes d'évaluer chaque item afin d'estimer si les apprenants des différents GPL du niveau scolaire concerné répondraient à chaque item correctement (oui ou non) (remarque : la **Figure II** n'est pertinente que lors de la définition de trois critères. Lorsqu'un critère de référence est défini, les animateurs peuvent simplifier ce graphique pour afficher uniquement les niveaux JPA et Dépasse le niveau (ADA), au lieu des niveaux AD). L'organigramme comporte trois points de décision qui doivent être pris en compte pour évaluer les éléments. Ces points de décision correspondent aux attentes des apprenants JPA, JA et JD décrites dans le GPF. Si un panéliste ne croit pas qu'un apprenant JD (un apprenant qui répond aux attentes décrites dans le descripteur Dépasse nettement les compétences minimales mondiales pour le niveau scolaire et le sous-concept) répondrait correctement à un élément d'une

¹⁹ Si, au cours de l'étape 1, 2 ou 3, le gouvernement décide qu'il souhaite uniquement fixer un critère de référence pour le niveau satisfaisant ou si l'auto-évaluation détermine que l'évaluation est trop courte pour prendre en compte trois critères de référence dans les trois principaux GPL, alors les panélistes nécessitent uniquement de conceptualiser les apprenants au niveau atteint ou JA.

évaluation, le panéliste encrclera DA, pour Dépasse nettement. En portant un jugement par oui ou par non aux trois points de décision, les panélistes doivent également prendre en compte les critères décrits cidessous qui décrivent le fait d'être « raisonnablement sûr » et d'estimer les performances des apprenants à chaque GPL/point de décision *lors d'une évaluation réelle dans la vie réelle, compte tenu des conditions d'évaluation et non la manière dont le GPF dit qu'ils devraient fonctionner. Cela signifie qu'ils prendront en compte les apprenants qui répondent aux attentes du GPL et du GPD appropriés et détermineront s'ils sont raisonnablement sûrs que ces apprenants répondraient correctement à l'item.*

Figure 11 : Processus d'évaluation des items pour la modification Angoff Oui-Non



En complétant l'étape 4, les panélistes doivent évaluer leurs items en fonction de quatre attentes, c'est-à-dire les chances que les apprenants identifiés/conceptualisés ayant des compétences minimales (tels que décrites dans le GPF) répondent correctement à chaque item :

- Probablement pas (« non »)
- Éventuellement possible (« non »)
- Raisonnablement sûr OU $\geq 67\%$ de chance OU deux apprenants sur trois (« oui »)
- Absolument positif (« oui »)

Pour répondre oui, les panélistes doivent être soit raisonnablement sûrs, soit absolument certains qu'un apprenant peu compétent répondrait correctement à l'item. Il convient également de demander aux panélistes de baser leurs évaluations sur « serait » plutôt que sur « devrait » afin de définir des attentes réalistes. Voici les définitions de « serait » et « devrait » :

- « Devrait » fait référence à une approche basée uniquement sur les performances, selon les GPD
- « Serait » est influencé par les contraintes d'évaluation, p. ex. la difficulté d'un item pour un apprenant particulier, les conditions de test, l'anxiété de l'apprenant et les erreurs aléatoires.

Remarque importante concernant les évaluations chronométrées : pendant le processus de notation, les panélistes travaillant avec une évaluation chronométrée devront suivre deux étapes :

1. Réfléchir au nombre d'items qu'un apprenant tenterait dans le temps imparti
2. Déterminer ensuite si l'apprenant aurait répondu correctement à chaque item (en suivant les étapes typiques de la tâche 3 décrites dans la **figure 11**).

Remarque importante pour les évaluations en lecture : lorsque les panélistes se demandent si des apprenants peu compétents répondraient correctement à un item, ils doivent également prendre en compte le niveau scolaire du mot ou le passage des références à l’item. Pour les mots, cette considération doit être basée sur les attentes des pays quant aux mots qui doivent être enseignés à un niveau scolaire spécifique, compte tenu de toutes les différences linguistiques entre les pays. Pour les passages, les panélistes devront prendre en compte les critères permettant de déterminer le niveau de qualité/la complexité du texte d’un passage, inclus dans les annexes A et C du GPF de lecture. Les détails sur la façon dont les panélistes devraient envisager d’évaluer les items en fonction de leur évaluation du niveau scolaire/de la complexité du texte d’un passage sont inclus dans la **figure 12**.

Figure 12 : Niveau scolaire/complexité du texte des passages de lecture

Présentation – Les GPD du GPF de lecture reposent largement sur l’hypothèse que l’évaluation liée comprend des mots et des passages adaptés au niveau scolaire. Cependant, ce n’est pas toujours le cas. Certaines évaluations incluent volontairement des passages de plusieurs niveaux scolaires afin que les résultats puissent aider les éducateurs à comprendre à quel niveau scolaire les apprenants réussissent. D’autres évaluations sont utilisées pour plusieurs niveaux scolaires d’apprenants afin d’examiner les progrès d’une année à l’autre. En outre, comme indiqué ci-dessus, les évaluations diffèrent considérablement selon leur niveau de difficulté. Pour cette raison, il est essentiel que les panélistes travaillant à relier les évaluations en lecture travaillent à déterminer le niveau scolaire des mots/passages de l’évaluation. Pour les mots, les panélistes auront besoin d’informations sur les mots enseignés dans le niveau scolaire concerné dans ce pays – probablement tirées du contenu national ou des normes de performance. Pour les passages, les panélistes doivent utiliser les annexes du GPF pour déterminer la complexité.

Détermination du niveau scolaire/de la complexité du texte – Pour les passages lus ou signés pour les apprenants (ceux qui correspondent au domaine Compréhension de la langue parlée ou signée), les panélistes doivent examiner les critères inclus dans l’annexe A du GPF de lecture. Pour les passages décodés par les apprenants (ceux qui correspondent aux domaines de décodage et/ou de compréhension écrite), les panélistes doivent examiner les critères inclus dans l’annexe B du GPF de lecture.

Lorsque le niveau scolaire du mot/du passage est approprié : si les panélistes estiment que le niveau scolaire du mot/du passage est correctement aligné sur le niveau scolaire pour lequel l’évaluation est liée, ils peuvent interpréter les GPD exactement tels qu’ils sont rédigés.

Lorsque le niveau scolaire du mot/du passage est trop faible : si les panélistes estiment que le niveau scolaire du mot/du passage est trop faible ou trop facile pour le niveau scolaire pour lequel l’évaluation est liée, ils doivent supposer qu’un apprenant peu compétent pourrait être en mesure de faire plus que ce qui est répertorié dans le >GPD du niveau de performance approprié. Dans quelle mesure cela dépend-il de la facilité du mot/du passage (p. ex., s’agit-il d’un niveau inférieur ou de deux ou trois niveaux inférieurs ?).

Lorsque le niveau scolaire du mot/du passage est trop élevé : si les panélistes estiment que le niveau scolaire du mot/du passage est trop élevé/difficile pour le niveau scolaire pour lequel l’évaluation est liée, ils doivent supposer qu’un apprenant peu compétent ne sera probablement pas en mesure de faire tout ce qui est indiqué dans le GPD du niveau de performance approprié. Combien moins dépend de la facilité du mot/passage.

Les panélistes doivent passer par deux séries d’évaluation sur deux jours différents, avec une discussion approfondie entre les deux séries. La littérature suggère que le fait de demander aux panélistes d’évaluer les éléments deux fois, au cours de deux séries distinctes, contribue à améliorer la qualité des notes ainsi que l’erreur type des critères de référence (SE) et la fiabilité inter-évaluateurs (voir **l’annexe J** pour plus de détails sur la façon de calculer ces valeurs et le **Chapitre V** pour plus de détails sur quand/pourquoi ceux-ci sont calculés), qui doivent être pris en compte dans le cadre du processus d’auto-évaluation à la fin de l’atelier pour déterminer si les résultats de l’atelier consacré à l’établissement de liens entre les politiques répondent aux exigences de fiabilité et de validité pour être accepté par l’ISU et d’autres donateurs pour les rapports mondiaux.

Au cours de la discussion qui a lieu entre les évaluations des séries 1 et 2, les animateurs doivent présenter aux panélistes :

- **un résumé de leurs notes** ainsi que la façon dont leurs notes individuelles se comparent à celles des autres panélistes. Cela devrait également guider les panélistes dans les discussions sur les points pour lesquels il y avait un désaccord considérable dans les notes oui/non.

- **des informations sur la difficulté des items** (des conseils sur la manière de générer ces données sont inclus dans le **chapitre V**), qui aident les panélistes à examiner leurs propres décisions concernant la difficulté des items.
- **des données d'impact** sur le pourcentage d'apprenants qui tomberaient dans chacun des GPL sur la base de l'itération la plus récente de l'évaluation (des conseils sur la manière de générer ces données sont inclus au **chapitre V**), ce qui aide les panélistes à avoir une idée de l'impact de leur notations et références.

Les panélistes doivent enregistrer leurs réponses à chaque série sur le même formulaire d'évaluation des items. Un exemple de formulaire – comportant six items – est présenté dans le **tableau 7**.

Tableau 7 : Formulaire d'évaluation des items à utiliser avec la modification Angoff oui-non

Item n°	Prédictions individuelles et indépendantes de la série 1				Prédictions individuelles et indépendantes de la série 2			
	JPA	JA	JD	DA	JPA	JA	JD	DA
1	JPA	JA	JD	DA	JPA	JA	JD	DA
2	JPA	JA	JD	DA	JPA	JA	JD	DA
3	JPA	JA	JD	DA	JPA	JA	JD	DA
4	JPA	JA	JD	DA	JPA	JA	JD	DA
5	JPA	JA	JD	DA	JPA	JA	JD	DA
6	JPA	JA	JD	DA	JPA	JA	JD	DA

Les panélistes doivent soumettre leurs formulaires aux animateurs à la fin de chaque série, et les animateurs synthétiseront le nombre de réponses oui par GPL pour obtenir un critère de référence individuel pour chaque panéliste. Les animateurs doivent ensuite faire la moyenne des critères de référence de chaque panéliste pour déterminer les critères de référence recommandés par le panel. Les alinéas ci-dessous montrent comment les notes des panélistes sont utilisées pour créer des références, à la fois pour chaque panéliste et pour l'ensemble du panel.

- Calculez les totaux pour les critères initiaux et finaux pour chaque panéliste :
 - Partiellement atteint = Total de chaque « oui » dans la colonne JPA du formulaire de notation
 - Atteint = Total de chaque « oui » dans les colonnes JPA et JA du formulaire de notation
 - Dépasse = Total des « oui » dans les colonnes JPA, JA et JD du formulaire de notation
- Calculez les moyennes des références mondiales initiales pour le panel :
 - Partiellement atteint = Moyenne des références partiellement atteintes pour tous les panélistes
 - Atteint = Moyenne des critères « atteint » pour tous les panélistes
 - Dépasse = Moyenne des critères « dépasse » pour tous les panélistes

Étant donné que les critères initiaux et finaux du panel sont calculés en prenant les moyennes des repères des panélistes, les repères auront presque toujours des valeurs fractionnaires, c'est-à-dire pas des nombres entiers. Lorsque cela se produit, les **critères de référence doivent toujours être arrondis** au point inférieur, même si cela va à l'encontre des règles mathématiques d'arrondi typiques. La raison en est que les critères désignent des niveaux de compétence minimaux et que l'avantage doit être accordé à l'apprenant (suivant le principe « ne pas nuire »).

Le calcul des critères finaux et la présentation des résultats par les animateurs principaux et l'analyste de données complètent l'atelier consacré à l'établissement de liens entre les politiques. Les détails du calcul des critères de référence sont inclus à **l'annexe T**. Les détails de la préparation de l'atelier sont

présentés au **chapitre IV** ci-dessous, et les notes de l'animateur pour la mise en œuvre de cette méthodologie dans un atelier en personne ou à distance sont incluses au **chapitre V**.

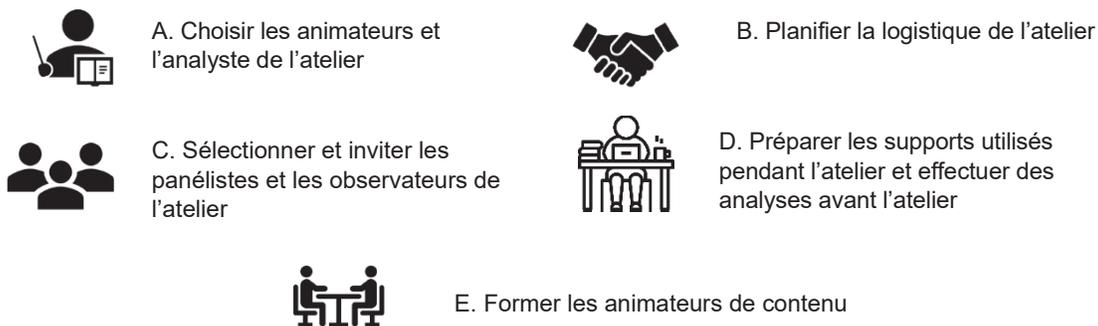
CHAPITRE IV

CHAPITRE IV. PRÉPARATION DE L'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Les responsables gouvernementaux/agents des agences d'évaluation et les représentants des donateurs, le cas échéant, auraient dû se rencontrer au cours de l'étape I : engagement initial pour parvenir à un accord sur l'opportunité d'établir un lien entre les politiques pour les rapports mondiaux et sur l'évaluation ou les évaluations qu'ils lieront aux normes mondiales à travers ce processus. Les ressources pour l'étape I sont liées dans le **tableau 3**. L'un des objectifs clés de la première étape est de garantir l'adhésion et l'appropriation du processus par le gouvernement ainsi que son engagement tout au long de la planification et de la préparation – avec l'intention que si le gouvernement ne met pas en œuvre l'atelier de lui-même, après l'atelier, il devrait avoir la capacité de refaire un atelier similaire pour fixer des critères supplémentaires sur différentes évaluations dans les années à venir si nécessaire.

Au cours de cette étape (Étape 4 : préparation de l'atelier d'établissement de liens entre les politiques), l'équipe du projet – composée de l'équipe de animateurs et de logisticiens du gouvernement ou des partenaires désignés pour diriger l'atelier – doit réaliser les cinq activités présentées dans la **figure 13**. Une liste de contrôle détaillée des préparatifs techniques et logistiques utilisée par l'équipe du projet, en collaboration avec les responsables gouvernementaux et les représentants des donateurs, se trouve à l'**annexe D**.

Figure 13 : Activités de préparation à l'atelier consacré à l'établissement de liens entre les politiques



A. CHOISIR LES ANIMATEURS ET LES ANALYSTE DE L'ATELIER

L'équipe du projet sélectionnera des animateurs et un analyste de données pour l'atelier sur la base de ces critères :

Animateur(s) principal(ux) – Responsable(s) de diriger l'atelier en s'assurant que les panélistes comprennent la méthode d'établissement de liens entre les politiques et ce qui est attendu. Ils doivent posséder une expertise en matière de liaison et d'analyse comparative des politiques, de solides compétences organisationnelles, d'excellentes compétences en présentation et une expérience avec des éducateurs allant des enseignants aux décideurs politiques. Ils doivent être conscients des défis liés au processus de liaison politique et des mesures correctives qui peuvent être prises pour relever ces défis.

Animateurs de contenu – Chargés d'aider les panélistes à interpréter et à comprendre le GPF et le contenu de l'évaluation, sur la base d'une compréhension de la langue et du contexte locaux. Les animateurs sont présents pour chaque évaluation, c'est-à-dire par matière, niveau et langue. Ils doivent

être capables d'apprendre rapidement puisqu'ils n'ont généralement pas d'expérience préalable en matière de liaison ou d'analyse comparative des politiques. Les animateurs de contenu doivent avoir une expérience dans les théories et techniques de mesure pédagogique, des compétences en animation de groupe et une expérience dans le domaine du contenu (lecture et/ou mathématiques) et le contexte. Ils doivent comprendre les normes du programme et du contenu, ainsi que la manière dont elles sont mises en œuvre par les enseignants en classe dans le contexte dans lequel les évaluations ont été mises en œuvre. Ils doivent maîtriser la langue de l'évaluation.

Analyste de données – Responsable de l'analyse des données de l'atelier et de l'organisation des informations à présenter aux panélistes. L'analyste pourrait être l'un des principaux animateurs possédant les compétences requises, si cette personne dispose de suffisamment de temps pendant l'atelier, bien qu'il soit recommandé de faire appel à un analyste de données dédié. Ce rôle nécessite une formation en statistiques, des compétences en informatique et en visualisation de données, ainsi que des compétences en logiciels (c-à-d. Excel ou Google Sheets pour les données de l'atelier ainsi que des logiciels statistiques, tels que Stata, SPSS ou R pour les données).

Il convient de noter qu'il est recommandé que les efforts de recrutement couvrent également un **coordinateur national d'atelier** et un **logisticien national**. Notez également que des animateurs peuvent également être sélectionnés à l'étape 1 pour aider à coordonner la collecte des documents par le gouvernement/l'agence d'évaluation à l'étape 2.

B. PLANIFIER LA LOGISTIQUE DE L'ATELIER

UTILISER LES ANNEXE D, ANNEXE E ET ANNEXE F

Trois options principales permettent d'organiser un atelier consacré à l'établissement de liens entre les politiques. Ceux-ci sont présentés ci-dessous avec leurs avantages et inconvénients. Il est recommandé, dans la mesure du possible, d'organiser des ateliers d'établissement de liens entre les politiques avec les animateurs et les panélistes réunis en présentiel, bien que des ateliers ont été organisés avec succès en utilisant des approches à distance et hybrides.

Tableau 8 : Options pour organiser un atelier consacré à l'établissement de liens entre les politiques

Option	Description	Avantages	Inconvénients
En présentiel	Les animateurs et les panélistes sont présents sur le même lieu pour l'atelier	<ul style="list-style-type: none"> • Offre une plus grande opportunité de discussion informelle entre les panélistes • Permet aux animateurs de fournir un soutien plus ciblé aux panélistes en fonction des besoins 	<ul style="list-style-type: none"> • Coût, susceptible de restreindre la représentativité géographique des panélistes • Peut être affecté par des restrictions de déplacement
À distance	Les animateurs et les panélistes assistent à l'atelier depuis leur propre domicile/bureau	<ul style="list-style-type: none"> • Permet une plus grande représentation géographique (voir toutefois les inconvénients associés liés aux implications technologiques) • Coût 	<ul style="list-style-type: none"> • Nécessite une gestion minutieuse pour garantir des discussions et un soutien de haute qualité aux panélistes • Peut avoir des implications technologiques, en particulier dans les zones où la connectivité est faible

Hybride	Les panélistes et les animateurs de contenu se réunissent en personne, dans le pays (en tant que groupe ou en tant que groupes régionaux) et les animateurs principaux assistent à distance (en supposant que les animateurs principaux sont	<ul style="list-style-type: none"> • Offre une plus grande opportunité de discussion informelle entre les panélistes • Permet aux animateurs de contenu d'aider les panélistes 	<ul style="list-style-type: none"> • Nécessite une gestion prudente par les animateurs principaux pour garantir des discussions et un soutien de haute qualité aux panélistes • Peut être affecté par des restrictions de déplacement
---------	--	--	---

L'équipe du projet doit travailler avec les parties prenantes gouvernementales et partenaires concernés pour sélectionner l'option appropriée en fonction du contexte, de la sécurité des participants et du budget. S'il est possible qu'au moins certains participants assistent en présentiel à l'atelier, l'équipe du projet devra travailler avec le gouvernement pour sélectionner un lieu approprié pour cette activité. S'il n'est pas possible de se réunir en présentiel, l'équipe du projet et le gouvernement doivent convenir d'une plateforme numérique fonctionnelle (en s'assurant que les licences adéquates sont achetées pour permettre l'accès à toutes les fonctionnalités pertinentes de la plateforme). Ils doivent également convenir et planifier d'autres aspects logistiques, par exemple si l'interprétation de l'atelier et/ou la traduction du matériel est nécessaire ; s'ils couvriront les frais de transport des panélistes, d'hôtel et les indemnités journalières ou les cartes de téléphone/Internet ; s'ils prennent en charge le repas pendant l'atelier ; s'ils enverront à l'avance l'évaluation ou un échantillon de celle-ci aux panélistes ; etc. Une liste de contrôle pour la préparation de l'atelier et un planificateur d'activités de l'atelier sont fournis à **l'annexe D** et à **l'annexe E**. Un modèle d'estimation budgétaire est fourni à **l'annexe F** pour aider les pays à estimer les coûts de l'atelier.

Outre la plateforme numérique sur laquelle héberger un atelier à distance ou hybride, l'équipe du projet devra également tenir compte des items suivants :

- **La connectivité** – Cela peut nécessiter l'achat de cartes de données internet pour s'assurer que les panélistes disposent de suffisamment de données pour se connecter aux sessions de l'atelier en bénéficiant d'une connexion suffisamment stable.
- **Le matériel (des panélistes)** – Il est fortement recommandé que les panélistes assistent à l'atelier avec un ordinateur ou un ordinateur portable équipé d'un écran suffisamment grand. Cela leur permettra de suivre plus clairement les présentations et tout autre matériel partagé à l'écran. Lorsque cela n'est pas possible et que les panélistes y assisteront par téléphone mobile, il sera essentiel de fournir des versions papier de tous les documents, notamment les diapositives de présentation.
- **Le matériel (dans la salle)** – Dans un atelier hybride, il est important de s'assurer qu'un écran est disponible afin que tous les panélistes puissent voir les animateurs principaux, en étant équipés de caméras et de microphones appropriés afin que les animateurs principaux puissent voir et entendre ce qui se passe dans la salle de l'atelier.
- **L'application de messagerie** – Lors d'ateliers à distance et hybrides, il est important que les panélistes puissent envoyer des messages aux animateurs. Une application de messagerie sécurisée (p. ex, WhatsApp, Telegram) doit être choisie à cette fin.

Enfin, outre la logistique générale, au cours de cette activité, l'équipe du projet doit s'entendre avec le(s) gouvernement(s) sur les moyens de poursuivre l'engagement avec le(s) gouvernement(s) du pays/l'agence d'évaluation qui a commencé avant l'atelier (durant l'étape 1). Cet engagement doit idéalement se poursuivre tout au long de l'atelier et après sa conclusion. L'objectif, avec l'engagement du gouvernement national/de l'agence d'évaluation, est de donner activement aux principaux représentants un rôle dans la préparation et la conduite de l'atelier, ce qui renforcera les capacités et permettra aux gouvernements et aux agences d'évaluation d'organiser de futurs ateliers si nécessaire.

C. SÉLECTIONNER ET INVITER LES PANÉLISTES DE L'ATELIER

Sélectionner les panélistes

SE REPORTER À L'ANNEXE M

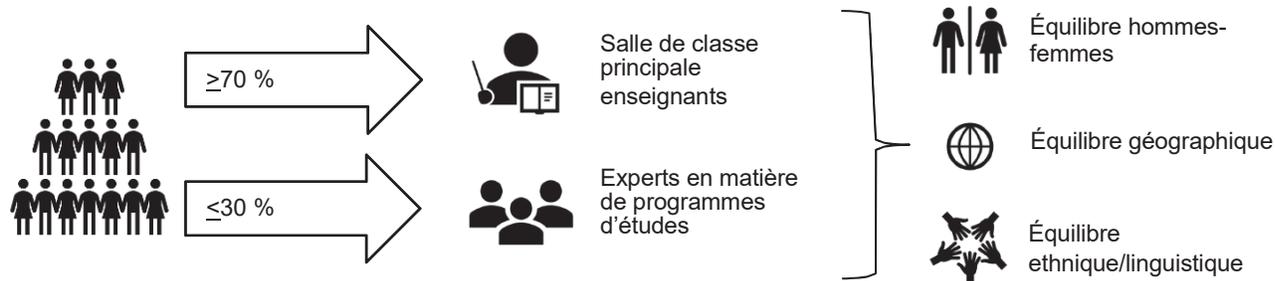
Les panélistes jouent un rôle clé dans l'atelier, car ce sont eux qui porteront réellement des jugements sur le lien entre l'évaluation ou les évaluations et le GPF, puis établiront des critères de référence pour l'évaluation ou les évaluations sur la base de ce lien. L'équipe du projet doit prévoir des panels distincts pour chaque niveau, matière et langue d'évaluation utilisés pour l'établissement de liens entre les politiques. Si plusieurs évaluations sont incluses dans un seul atelier, p. ex., lecture de troisième année et mathématiques de troisième année, des séances plénières seront consacrées à la formation, la discussion et la présentation, mais chaque panel aura des activités de groupe distinctes pour vérifier l'alignement avec le GPF, faire correspondre les items avec les GPL et les GPD, et définir les critères de référence.

Lors de la sélection d'un ou plusieurs panels pour un atelier consacré à l'établissement de liens entre les politiques, le nombre de panélistes doit être suffisamment grand et représentatif. Il s'agit de fournir une assurance raisonnable que les critères 1) seront réalistes, réalisables et impartiaux et 2) ne varieront pas beaucoup si le processus était répété avec différents panélistes. Les panélistes doivent posséder de solides connaissances du contenu et des compétences pédagogiques (lecture ou mathématiques). Ils doivent être qualifiés pour porter les jugements qui leur sont demandés afin d'établir des critères de référence. Les panélistes doivent être perçus comme des experts dans leur domaine au sein de leur système éducatif afin de favoriser la confiance des gouvernements hôtes dans leurs décisions.

Pour chaque évaluation, un groupe de 15 panélistes est un minimum et 20 panélistes un maximum. Un groupe de cette taille garantira que le processus obtiendra un résultat reproductible mais sera également pratique et gérable.²⁰ Comme le montre la **Figure 14**, le panel doit être composé d'au moins 70 pour cent d'enseignants de classe principale et jusqu'à 30 pour cent de non-enseignants, de préférence des experts en matière de programmes d'études.

²⁰ Se référer à Livingston et Zieky, 1982 ; Norcini, Shea et Grasso, 1991; Mehrens et Popham, 1992; Hurtz & Hertz, 1999 pour la littérature en ce qui concerne la taille du panel ainsi que les caractéristiques et qualifications des panélistes.

Figure 14 : composition des panélistes



La composition type d'un panel est de 12 enseignants et de 3 experts en matière de programme d'études. Les qualifications des panélistes comprennent les items suivants :

- Au moins cinq ans d'enseignement au niveau scolaire concerné ou adjacent (enseignants)
- Au moins cinq ans d'expérience dans l'enseignement (experts en programme d'études)
- Excellentes compétences dans la matière (lecture ou mathématiques)
- Compétences natives dans la langue d'enseignement et d'évaluation
- Expérience avec divers apprenants à différents niveaux de compétence
- Connaissance du système d'enseignement, y compris du matériel
- Certification et permis d'enseignement collégial et/ou universitaire

Au-delà des qualifications, la représentativité des panels doit être assurée à travers les critères suivants :

- **Représentation des sexes** – Les panélistes doivent être sélectionnés pour garantir un équilibre entre les sexes proportionnel à la profession enseignante dans le pays, tant pour le personnel enseignant que les non-enseignants.
- **Représentation géographique** – Les panélistes doivent être sélectionnés de manière à garantir la représentation des régions, des provinces et/ou des États des évaluations.
- **Représentation ethnique et/ou linguistique** – Les panélistes doivent avoir une diversité qui reflète la population ainsi que la langue des évaluations.
- **Autre représentation** – En fonction de sa pertinence par rapport au contexte et aux populations d'apprenants spécifiques pour lesquelles les résultats seront communiqués, la composition des enseignants et des non-enseignants peut devoir refléter également d'autres caractéristiques. Ces caractéristiques pourraient inclure les items suivants : affectation dans des écoles privées et publiques, expérience avec des apprenants en situation de handicap, expérience dans des programmes d'apprentissage accéléré et emplacement dans des environnements de crise et de conflit.
- **Représentation pour les évaluations multinationales** – Lorsque l'atelier consacré à l'établissement de liens entre les politiques cherche à relier les évaluations régionales ou internationales au GPF, il est important que les panélistes représentent plusieurs pays ou que des ateliers distincts soient organisés pour chaque pays, puis que les résultats soient comparés pour déterminer les références finales. Les animateurs doivent contacter le comité d'examen 4.1.1 pour plus de détails sur la représentation appropriée dans les évaluations régionales/internationales.

L'équipe du projet doit collaborer avec le gouvernement, l'agence donatrice, le(s) partenaire(s) de mise en œuvre et/ou d'autres parties prenantes pour déterminer la manière la plus appropriée de recruter des panélistes. Cela peut se faire par le biais de nominations par le ministère de l'Éducation, l'unité

d'évaluation ou un autre organisme gouvernemental. Le gouvernement, le donateur, le partenaire et les animateurs doivent discuter de la manière d'appliquer les critères dans leur contexte. Il est important que les différentes parties conviennent d'exigences minimales en matière de qualifications et de critères de représentativité. Les données démographiques finales des panélistes doivent être collectées, regroupées et soumises avec les résultats de l'atelier en utilisant le formulaire inclus à **l'annexe M**. Remarque : les animateurs peuvent souhaiter envoyer ce formulaire par voie électronique aux panélistes invités avant l'atelier pour confirmer la représentativité du panel, ou les animateurs peuvent l'imprimer et le récupérer auprès des panélistes pendant l'atelier. Ce formulaire fournira à l'équipe de projet suffisamment de données pour déterminer dans quelle mesure les panélistes répondent aux critères dans le cadre de l'auto-évaluation post-atelier.

Une liste des participants et leurs coordonnées doivent être disponibles suffisamment à l'avance pour que les activités de préparation puissent être correctement menées.

Inviter des panélistes et des observateurs et mener l'activité préalable à l'atelier

UTILISER L'ANNEXE K ET L'ANNEXE L

Les panélistes doivent être invités bien avant l'atelier ; au moins trois à quatre semaines sont recommandées. **L'annexe K** et **l'annexe L** comprennent des projets de lettres d'invitation pour les observateurs (p. ex., les représentants du gouvernement/de l'agence d'évaluation, les donateurs, d'autres donateurs/partenaires internationaux ou locaux qui pourraient être intéressés à organiser un futur atelier d'établissement de liens entre les politiques ou à comprendre le processus à des fins plus générales) et panélistes respectivement. Les lettres d'invitation doivent inclure des informations de base sur l'atelier et la logistique, c.-à-d. les objectifs, les attentes, les dates, le transport, l'hébergement, les repas et les indemnités journalières.

Dans la mesure du possible, les invitations doivent inclure l'outil ou les outils d'évaluation complets qui seront liés aux normes mondiales avec des instructions sur la manière dont ils doivent être administrés aux apprenants avant l'atelier. Si le gouvernement rencontre des problèmes de sécurité liés à la publication de l'évaluation, un échantillon de questions d'évaluation peut être utilisé, comme décrit dans les alinéas suivants. Toutefois, il n'est pas préférable d'utiliser un échantillon d'évaluation en lieu et place de l'évaluation complète, car cela ne donnera pas aux panélistes un aperçu des critères de référence raisonnables. Voir la **figure 15** pour plus d'informations sur la sécurité de l'évaluation.

- Pour les évaluations chronométrées administrées individuellement, telles que les évaluations de lecture ou de mathématiques en début de scolarité (EGRA ou EGMA), les exemples d'évaluations comprendront des sous-tâches de lecture ou de mathématiques, selon le cas.
- Pour les évaluations non chronométrées administrées en groupe ou individuellement, telles que la plupart des évaluations basées sur le programme d'études (ACA), les exemples d'évaluations comprendront des items de lecture ou de mathématiques, selon le cas.

Au cours de l'atelier, les panélistes recevront une formation supplémentaire et une expérience pratique en matière d'administration et de notation des évaluations.

Compte tenu de la longueur du document, il peut également être utile que les panélistes reçoivent le GPF à ce stade. Cependant, l'équipe du projet devra décider de la meilleure façon d'introduire le GPF, car les panélistes peuvent faire preuve de confusion si la note de l'évaluation est différente de celle du GPF utilisé pour l'établissement de liens entre les politiques.

Figure 15 : considérations relatives à la sécurité de l'évaluation

Raisons de la sécurité de l'évaluation – Pour éviter que les enseignants n'enseignent en fonction des tests ou que les apprenants ne trichent aux tests, il est important de maintenir la sécurité des outils d'évaluation.

Quels tests doivent être sécurisés – La sécurité est la plus critique pour les CBA, en particulier pour ceux administrés à tous les apprenants d'un niveau particulier à l'échelle nationale. La sécurité des évaluations qui sont administrées uniquement à un échantillon d'apprenants et/ou qui changent régulièrement (par exemple chaque année) est moins importante. Cependant, les protocoles de sécurité doivent être laissés au gouvernement/à l'agence d'évaluation.

Protocoles de sécurité pour les ateliers consacrés à l'établissement de liens entre les politiques – Les protocoles de sécurité d'évaluation varient en fonction des préférences du gouvernement et/ou de l'agence d'évaluation. Cependant, les protocoles de sécurité suivants sont souvent utilisés avec les CBA :

- **Activité préalable à l'atelier** – Si l'évaluation est mise en œuvre avec un recensement des apprenants ou n'est pas modifiée régulièrement, le gouvernement/l'agence d'évaluation souhaitera peut-être envoyer uniquement un échantillon de questions de l'évaluation ou un échantillon d'items d'évaluation similaires.
- **Protocoles d'atelier** – Les évaluations ne peuvent pas être incluses dans les dossiers des panélistes, mais peuvent plutôt être distribuées avec les numéros d'identification des panélistes (voir la section D de ce chapitre pour en savoir plus sur les numéros d'identification des panélistes et la préparation des dossiers) indiqués en haut au début de chaque journée ou pour chaque activité dans laquelle l'évaluation est nécessaire, puis collectée à la fin de la journée ou de l'activité.

D. DOCUMENTS ET ANALYSES

UTILISER LES ANNEXE A, ANNEXE G, ANNEXE H, ANNEXE I, ANNEXE M, ANNEXE N, ANNEXE O, ANNEXE P, ANNEXE Q, ANNEXE R ET ANNEXE S

Tous les documents et analyses nécessaires pour l'atelier sont répertoriés ci-dessous dans une série de trois listes, organisées par documents qui doivent être obtenus auprès du gouvernement ou de l'agence d'évaluation régionale/internationale, analyses qui doivent être menées à l'aide de ces documents avant l'atelier, et les documents qui doivent être créés/adaptés. L'utilisation de chacun de ces documents dans l'atelier est également mentionnée dans les chapitres et sections suivants.

La plupart d'entre eux auraient dû être obtenus lors de la réalisation de l'activité d'auto-évaluation ; ainsi, si les animateurs ont participé à cette étape, ils devraient déjà avoir accès à tous sauf aux items étoilés ci-dessous (qu'ils devront demander).

Certains documents devront être traduits dans la langue dans laquelle l'atelier se déroulera pour les animateurs (les panélistes doivent parler couramment la langue de l'évaluation). Ceux-ci sont indiqués par un (†). La décision concernant les langues pour lesquelles fournir une traduction sera déterminée par le pays. Il se peut que la documentation ne soit pas traduite dans les premières langues pour tous les panélistes, en fonction du nombre de langues parlées. En revanche, tous les panélistes doivent maîtriser suffisamment au moins une des langues dans lesquelles les documents sont traduits. Certains documents doivent être fournis aux panélistes sous forme papier. Ceux-ci sont indiqués par un (‡)

Afin d'obtenir ces documents, les gouvernements peuvent exiger l'élaboration et la signature d'un accord de confidentialité.

Documents à obtenir

- Caractéristiques de l'évaluation
- Tous les outils d'évaluation utilisés dans l'évaluation (†) (‡)
- Ensemble complet de fichiers de données d'évaluation
- Corrigés et grilles de notation

- Normes nationales sur la fluidité/le rythme de décodage et le texte de niveau scolaire (si définies et si les pays associent une évaluation de la lecture)*
- Rapport technique, notamment les résultats de la mise en œuvre la plus récente de l'évaluation
- Échantillon(s) d'évaluation(s), créé(s) sur la base de l'évaluation complète (si nécessaire à des fins de sécurité, comme décrit dans la **Section C**)* (‡)

La plupart de ces documents/données serviront à l'analyse qui doit avoir lieu avant l'atelier, qui est décrite en détail ci-dessous.

Les analyses qui devraient être effectuées

Les animateurs doivent calculer/préparer les informations sur les items suivants avant l'atelier à l'aide de l'évaluation, du fichier de données, du corrigé et des grilles de notation (le cas échéant) :

- **Difficulté des items** – Consulter l'**annexe N** pour plus de détails sur la façon de calculer ces statistiques à l'aide des données des résultats d'évaluation les plus récents.
- **Distributions des données** – Consulter l'**annexe O** pour plus de détails sur la façon de préparer ces données. Les distributions de données montreront le nombre et le pourcentage d'apprenants qui ont passé l'évaluation et qui ont obtenu tous les scores possibles à l'évaluation. Bien que ces données puissent être préparées avant l'atelier, elles ne sont pas nécessaires avant le jour 4, lorsqu'elles constitueront la base de l'analyse des informations d'impact entre les séries 1 et 2 de notation Angoff (quel pourcentage d'apprenants satisferaient chacun des GPL sur la base des notes/critères initiaux du panéliste et les données de l'itération la plus récente de l'évaluation).

Cette analyse éclairera la série 2 des notations Angoff menées lors de la tâche 3.

Documents et données à créer/adapter

L'équipe du projet/les animateurs de l'atelier doivent créer (ou adapter à partir des modèles/exemples fournis dans ce référentiel) les documents suivants :

- **Programme de l'atelier** – Un modèle est inclus à l'**annexe P**, qui devra être adapté comme décrit ci-dessous. (†)
- **Identifiants des panélistes**–Ils doivent être attribués le premier jour de l'atelier et doivent être confidentiels entre le panéliste et l'équipe de projet.
- **Feuille de présence journalière** – Elle doit être créée et suivie pendant l'atelier pour garantir que chaque panéliste a reçu toute la formation nécessaire.
- **Informations démographiques sur les panélistes** – Le formulaire est inclus à l'**annexe M** mais peut devoir être mis à jour en fonction des critères de représentativité des panélistes.
- **GPD de niveau/matière pertinent** – Pour les ateliers liés à l'ODD 4.1.1(a), les panélistes doivent recevoir les niveaux 1, 2 et 3 du GPF. Pour les ateliers liés à l'ODD 4.1.1(b), les panélistes doivent recevoir les niveaux 4, 5 et 6 du GPF. Pour les ateliers liés à l'ODD 4.1.1(c), les panélistes doivent recevoir les niveaux 7, 8 et 9 du GPF. (Les animateurs devront réduire le GPF aux seuls GPL conformes si les critères de référence ne sont fixés que pour un seul GPL.) De plus, si l'évaluation est une évaluation en lecture, les annexes pertinentes doivent également être incluses dans le dossier afin que les panélistes disposent de critères pour évaluer le niveau scolaire d'un passage de lecture, ainsi que des exemples d'items pour les niveaux scolaires concernés. (†) (‡)

- **Diapositives de facilitation** (atelier) – Des détails sur la façon de localiser les modèles de diapositives sont inclus à **l'annexe G** pour les évaluations chronométrées et non chronométrées, mais les animateurs devront les adapter ; les instructions sur la façon de procéder sont incluses dans le modèle. (†) (‡)
- **Diapositives de facilitation** (formation de animateur de contenu) – Des détails sur les diapositives à utiliser parmi les principales diapositives de facilitation sont inclus à **l'annexe S**. (†)
- **Formulaires d'évaluation de l'alignement et formulaires d'évaluation des items** – Des détails sur la façon de créer ces formulaires et des exemples sont inclus à **l'annexe H** pour le formulaire d'alignement et **l'annexe I** pour le formulaire d'évaluation des items.
- **Formulaires d'évaluation d'atelier** – Une ébauche est incluse à **l'annexe R**. L'équipe du projet souhaitera peut-être ajouter des questions au formulaire et/ou le transformer en un formulaire d'évaluation quotidienne. (†)
- **Données de rétroaction sur l'atelier** (Notez que celles-ci ne peuvent être créées qu'après les évaluations des panélistes de l'étape 1, puis les notes de la série 2 ; les instructions sur la façon de générer ces données sont incluses à **l'annexe O**).

Les détails sur la manière de créer ou d'adapter ces supports et données, à l'exception de la feuille de présence, qui doit être intuitive, sont inclus ci-dessous :

PROGRAMME DE L'ATELIER

Le programme de l'atelier doit rester similaire, avec les mêmes tâches, que l'atelier prévu soit en présentiel, hybride ou à distance (le programme à distance devrait toutefois permettre des journées raccourcies et plus de pauses sur une plus longue période en raison de la fatigue provoquée par les réunions en ligne). Quel que soit le type d'atelier, une certaine flexibilité sera toujours nécessaire si les tâches prennent plus de temps que prévu ou si les panélistes terminent plus rapidement. Certaines activités devraient avoir lieu à la fin d'une journée/session, afin de permettre cette flexibilité et de mobiliser des animateurs pour préparer la session à venir.

Le **tableau 9** donne un aperçu des tâches de l'atelier, avec plus de détails donnés au **chapitre V** – Remarque : les activités au sein d'une tâche peuvent se dérouler sur des jours consécutifs pour un atelier en présentiel ou sur plusieurs sessions pour des ateliers en distanciel. Le modèle de programme de l'atelier (**annexe P**) fournit les répartitions de temps et les exigences de facilitation pour les activités de chaque tâche. Des exemples de programmes des ateliers en présentiel et à distance sont fournis à **l'annexe Q**.

La structure des tâches doit rester constante pour tous les ateliers, même si de légères modifications peuvent être nécessaires dans la répartition du temps en fonction de la logistique et d'autres problèmes spécifiques au contexte. Les animateurs doivent revoir le programme, ajuster les dates, ajuster les heures des pauses (en fonction des normes locales), ajouter tous les discours nécessaires des représentants du gouvernement, des agents des agences d'évaluation, des donateurs, etc., puis envoyer au gouvernement/à l'agence d'évaluation et à ses partenaires leur examen avant finalisation. Il est recommandé que les ateliers en présentiel se déroulent pendant 8 heures par jour, sur 6 jours. Il est recommandé pour les ateliers à distance que les sessions se déroulent en sessions d'environ 2 à 4 heures, réparties sur une période plus longue de deux semaines à un mois. Cette dernière période est destinée à permettre aux panélistes d'examiner le GPF et de s'entraîner à administrer l'évaluation avant l'atelier, comme recommandé dans la sous-section « Inviter des panélistes et l'activité préalable à l'atelier » ci-dessus.

Tableau 9 : brève description des séances de l'atelier

Tâches	Descriptions
Ouverture	Cette tâche accueille les panélistes à l'atelier et prévoit du temps pour les présentations et se déroulera généralement en une seule session. Les dignitaires du ou des pays hôtes, y compris le(s) gouvernement(s), l'agence d'évaluation (le cas échéant) et l'agence donatrice (le cas échéant), sont invités à prendre la parole à l'atelier. Le coordinateur de l'atelier examine la logistique. Les animateurs principaux présentent le programme, les objectifs et un résumé de haut niveau de la méthode (en reconnaissant que cela sera nouveau pour la plupart des panélistes et ne devrait pas entrer dans des détails qui pourraient prêter à confusion pour les panélistes s'ils étaient présentés trop tôt).
Familiarisation	Cette tâche a pour objectif de présenter et d'examiner attentivement le GPF et les outils d'évaluation avant les activités au cours desquelles ces documents seront utilisés. Pour cette tâche, deux sessions seront organisées. Elles pourront se tenir le même jour ou séparément (en effet, cette tâche peut avoir lieu avant l'atelier principal si le temps est limité ou pour donner plus de temps aux panélistes pour comprendre les documents). Pour la séance sur l'outil d'évaluation, les animateurs peuvent choisir de demander aux panélistes de s'administrer l'évaluation les uns aux autres pour s'entraîner (évaluations chronométrées) ou de passer l'évaluation eux-mêmes (évaluations non chronométrées). À la fin de cette tâche, les panélistes achèvent la première partie de l'évaluation.
Alignement (Tâche 1)	Les animateurs principaux forment les panélistes sur la tâche d'alignement. Les animateurs de contenu dirigent l'activité de la tâche 1 sur l'alignement des évaluations avec le GPF, qui est une activité individuelle et indépendante. Les animateurs principaux présentent les résultats de l'alignement. À la fin de cette tâche, les panélistes achèvent la partie suivante de l'évaluation.
Mise en correspondance (Tâche 2)	Les animateurs principaux forment les panélistes à la tâche de mise en correspondance. Les animateurs de contenu dirigent l'activité de la tâche 2 sur la mise en correspondance des évaluations avec les GPD/GPL, qui est une activité de groupe. À la fin de cette tâche, les panélistes achèvent la partie suivante de l'évaluation.
Analyse comparative (Tâche 3)	Les animateurs principaux présentent un aperçu de l'analyse comparative mondiale. Ils forment les panélistes à la méthode Angoff. Les animateurs de contenu dirigent la première activité de la tâche 3 avec la méthode Angoff. Les animateurs de contenu dirigent la deuxième activité de la tâche 3 avec la série 1 Angoff. Les animateurs principaux analysent les résultats de la série 1 (cette activité devra avoir lieu pendant la nuit/entre les sessions) Les animateurs principaux présentent les résultats de la série 1. Les animateurs de contenu dirigent la troisième activité de la tâche 3 avec la série 2 Angoff. Les animateurs principaux présentent les résultats de la série 2. À la fin de cette tâche, les panélistes achèvent la dernière partie de l'évaluation.
Clôture	Les dignitaires du/des pays hôte(s), y compris le(s) gouvernement(s), l'agence d'évaluation (le cas échéant) et l'agence donatrice (le cas échéant), sont invités à clôturer l'atelier.

IDENTIFIANTS DES PANÉLISTES

Les panélistes doivent se voir attribuer des identifiants uniques et confidentiels (entre l'équipe de projet et le panéliste) avant l'atelier. Ils les utiliseront pour s'identifier sur leurs formulaires d'évaluation afin que les animateurs puissent effectuer un suivi auprès des panélistes qui ne semblent pas comprendre les concepts et pour que les évaluations anonymes des panélistes (informations normatives) puissent être présentées aux panélistes entre les évaluations des séries 1 et 2 et après la série 2, comme décrit plus en détail ci-dessous. Chaque panéliste doit connaître son numéro d'identifiant. Il peut être écrit sur un morceau de papier dans leur dossier ou quelque part à l'intérieur du dossier.

FICHE DE PRÉSENCE JOURNALIÈRE

Il est important de noter les présences chaque jour à l'atelier afin que les animateurs sachent quels panélistes ont manqué des sessions et puissent assurer le suivi auprès d'eux, le cas échéant, en vue de s'assurer qu'ils comprennent ce qu'ils doivent faire.

INFORMATIONS DÉMOGRAPHIQUES DES PANÉLISTES

Il est important de recueillir toutes les informations incluses dans le formulaire de **l'annexe M** pour garantir que les panélistes sont représentatifs de la population évaluée. Ces informations doivent également être communiquées au comité d'examen 4.1.1, ainsi que des détails sur la population évaluée et les enseignants de cette population. Par exemple, le comité d'examen 4.1.1 voudra des détails sur le pourcentage d'enseignants hommes-femmes de niveau X dans le domaine d'évaluation afin de vérifier la représentation des sexes des enseignants. Le formulaire inclus à **l'annexe M** devra peut-être mis à jour afin qu'il pose les questions appropriées sur les données démographiques et géographiques. Par exemple, certains pays n'ont pas de régions ou de districts mais plutôt des États ou des municipalités. Le formulaire peut être envoyé aux panélistes avant l'atelier ou distribué et récupéré pendant l'atelier.

GPD DE NIVEAU/MATIÈRE PERTINENTS

Le GPF est disponible sur Edulinks et sur le site internet de l'ISU. Il n'est cependant pas nécessaire de présenter aux panélistes l'intégralité du GPF. Au lieu de cela, les animateurs peuvent créer une version modifiée qui ne comporte que les niveaux pertinents (2^{ème} année pour l'ODD 4.1.1(a), 5e année pour l'ODD 4.1.1(b) et 8e année pour l'ODD 4.1.1(c) et le note ci-dessous. Les animateurs guideront les panélistes dans un examen attentif de ces tableaux au cours de l'atelier.

Le tableau des connaissances ou compétences du GPF, le tableau 3 et le tableau 5, qui comprend les GPD pour chacun des GPL, sont les plus utiles pour les ateliers axés sur l'établissement de trois critères – un pour chacun des GPL. Les ateliers axés sur l'établissement d'un seul critère de référence doivent utiliser les tableaux 3 et 4 du GPF. Dans les deux cas, les panélistes auront recours au tableau 3 pour la tâche 1 (Alignement). En fonction du nombre de critères qui seront définis, ils utiliseront ensuite soit le tableau 4 (pour un critère) soit le tableau 5 (pour trois critères) pour la tâche 3 (Notation). Le tableau 1 du GPF définit chaque GPL et constitue une référence utile pour les panélistes s'ils ne se souviennent pas d'un GPL spécifique. Le tableau 2 illustre les domaines, les concepts et les sous-concepts à travers les niveaux scolaires et fournit un résumé utile pour les décideurs politiques et les panélistes.

Notez encore une fois que si l'évaluation est une évaluation en lecture, les annexes GPF pertinentes doivent également être incluses dans le dossier afin que les panélistes disposent de critères pour évaluer le niveau scolaire d'un passage en lecture pour le niveau scolaire lié ainsi qu'un niveau supérieur et un inférieur également à titre d'exemple d'items pour les niveaux scolaires concernés.

Les animateurs, en collaboration avec le gouvernement, doivent déterminer si les deux tableaux, au minimum, pourraient devoir être traduits si la langue de l'évaluation ne se fait pas en anglais (voir la **figure 16** pour plus de détails), mais les animateurs ne doivent pas apporter d'autres modifications au contenu ou langage du GPF.

Figure 16 : Traduction du GPF

Des cabinets de traduction ou des traducteurs individuels peuvent aider à la traduction, mais la traduction doit être dirigée par des experts en contenu. Il est essentiel que le sens de chaque terme soit traduit de manière complète et précise et que la traduction des exemples destinés à la lecture implique de modifier les exemples, si nécessaire, pour garantir qu'ils restent adaptés au niveau scolaire (puisque la longueur et la complexité des mots peuvent changer en traduction). L'équipe du projet doit également envisager une traduction vers l'anglais pour valider la traduction dans une autre langue.

Enfin, au fil du temps, les GPD (et même l'ensemble du GPF) seront traduits dans de nombreuses langues, dont certaines pourront être utilisées dans plusieurs pays qui utilisent les mêmes langues. Même avec ces traductions, chaque pays doit lire attentivement les GPD traduits et apporter toutes les modifications nécessaires en fonction de l'usage de la langue locale.

DIAPOSITIVES DE FACILITATION (ATELIER PRINCIPAL ET FORMATION DES ANIMATEURS DE CONTENU)

Les animateurs présenteront les diapositives pendant les jours 1 à 6 de l'atelier (pour les ateliers en personne) ou à travers une série de huit séances d'atelier (pour les ateliers à distance). Les diapositives sont incluses à **l'annexe G** et comprennent des détails sur : 1) le programme, les objectifs et la méthode, 2) comment présenter le GPF et l'évaluation, 3) l'alignement, 4) la mise en correspondance, 5) l'analyse comparative et 6) l'évaluation. Notez que les diapositives ont nécessité une certaine adaptation selon la nature de l'évaluation.

L'équipe du projet doit consulter le gouvernement et d'autres parties prenantes clés pour déterminer si les diapositives de facilitation doivent être traduites dans la langue de l'évaluation ou dans une autre langue internationale. Si les diapositives ne sont pas traduites dans les langues locales, les animateurs de contenu ou les traducteurs peuvent les interpréter selon les besoins.

Les diapositives pour la formation des animateurs de contenu (**Annexe S**) sont un sous-ensemble des diapositives de l'atelier principal.

FORMULAIRES D'ALIGNEMENT ET D'ÉVALUATION DES ITEMS

Il existe deux types de formulaires de notation. L'équipe du projet adaptera les formulaires pour qu'ils correspondent à l'outil d'évaluation et aux parties pertinentes du GPF.

- **Formulaires d'évaluation de l'alignement (Annexe H)** – Ceux-ci seront utilisés pour évaluer par les panélistes l'alignement entre les évaluations et le GPF.
- **Formulaires d'évaluation des items (Annexe I)** – Ceux-ci seront utilisés pour les notes des panélistes de chaque item d'évaluation par rapport aux GPL et GPD.

Les annexes comprennent des exemples de formulaires d'alignement et d'évaluation des items issus d'évaluations chronométrées et d'évaluations non chronométrées. Les formulaires devront être adaptés d'une évaluation à l'autre en fonction du format d'évaluation (p. ex, nombre de domaines et de concepts), du ou des types de questions (p. ex, choix multiple ou mot unique) et de la notation (p. ex., dichotomique ou polytomique). Le formulaire d'évaluation de l'alignement a été créé pour être utilisé facilement, mais l'équipe du projet souhaitera peut-être le mettre à jour pour le rendre plus dynamique, avec des menus déroulants et des totaux générés automatiquement. Plusieurs options et exemples de formulaires d'évaluation des éléments sont inclus à **l'annexe I** avec des détails sur la manière de choisir et d'adapter les formulaires.

FORMULAIRES D'ÉVALUATION DE L'ATELIER

Les panélistes doivent a minima remplir une évaluation à la fin de l'atelier. Cependant, idéalement, les panélistes doivent remplir la section appropriée du formulaire d'évaluation à la fin de chaque tâche, pour vérifier l'acquisition de connaissances, les domaines qui peuvent nécessiter davantage de clarté, les techniques de facilitation qui fonctionnent/ne fonctionnent pas, etc. **L'annexe R** comprend les questions d'évaluation minimales qui devraient être posées aux panélistes d'ici la fin de l'atelier. Il est conçu pour recueillir leurs points de vue sur le processus d'établissement de liens entre les politiques et soutenir l'auto-évaluation du **chapitre VI**. Le formulaire se compose d'échelles de type Likert et de questions ouvertes sur la satisfaction des panélistes à l'égard de l'orientation, de la formation et du processus. Les résultats fourniront la preuve de la confiance des panélistes dans leurs jugements et solliciteront des commentaires supplémentaires sur l'expérience d'établissement de liens entre les politiques.

Si l'équipe du projet effectue une évaluation quotidienne qui identifie les problèmes nécessitant un recyclage, elle doit administrer à nouveau l'évaluation après la formation supplémentaire pour confirmer que les panélistes sont désormais satisfaits. Cela sera nécessaire pour soutenir l'auto-évaluation abordée au **Chapitre VI**.

Si l'équipe du projet choisit de ne pas inclure d'évaluation quotidienne, les animateurs principaux et les animateurs de contenu devraient au minimum envisager de procéder à des vérifications orales avec les panélistes à la fin de chaque journée pour discuter des débats et des adaptations possibles, p. ex., une meilleure interprétation des présentations en langue locale, la nécessité de revoir les étapes d'une tâche, etc.

DONNÉES DE RÉTROACTION SUR L'ATELIER

Les données de rétroaction sur l'atelier comprennent des informations normatives sur les notes des panélistes et les données d'impact. (Ces analyses auront lieu pendant l'atelier, pas avant). Les instructions sur la façon de générer ces statistiques et ces tableaux de rétroaction sont incluses à **l'annexe O**. L'analyste de données devra calculer les statistiques, les graphiques et les diagrammes à l'aide des données d'évaluation des panélistes de la série 1. En tant que tel, cela devra être fait entre les sessions pertinentes de l'atelier. Le même processus devra être achevé après les notations de la série 2 et un temps suffisant doit être prévu au programme pour que cela ait lieu.

DOSSIERS D'ATELIER

Une fois tous les documents créés ou adaptés et les données générées, l'équipe du projet devra imprimer les documents suivants à inclure dans chacun des dossiers des panélistes (et postés ou livrés aux panélistes dans le cas d'ateliers à distance) :

- Programme
- Identifiant du panéliste (peut être écrit en petits chiffres à l'intérieur du dossier ou imprimé sur une feuille de papier incluse dans le dossier)
- Lexique des termes (peut être imprimé à partir de celui inclus au début de ce document)
- Liste des acronymes (peut être imprimée à partir de celle incluse au début de ce document)
- GPD de niveau/matière pertinents du GPF (2^{ème} année pour l'ODD 4.1.1(a), 5^{ème} année pour l'ODD 4.1.1(b) et 8^{ème} année pour l'ODD 4.1.1(c))

- Outil d'évaluation (ne doit être inclus que si les protocoles de sécurité de l'évaluation le permettent ; voir la **Figure 15** pour plus de détails sur la sécurité de l'évaluation)
- Diapositives (imprimées en format notes)
- Formulaire d'évaluation d'alignement
- Formulaire d'évaluation des items

E. FORMER LES ANIMATEURS DE CONTENU

Les animateurs principaux devront organiser une session de formation pour les animateurs de contenu, qui ne sont probablement pas familiers avec la méthodologie d'établissement de liens entre les politiques. Un modèle de diapositive de formation pour le animateur de contenu est disponible à l'**annexe S**. La formation doit inclure un aperçu du programme de l'atelier ; une discussion détaillée du GPF ; un examen de l'évaluation ou des évaluations ; et des exercices pratiques d'alignement, de mise en correspondance et d'analyse comparative. Ces exercices pratiques sont essentiels pour garantir l'engagement du animateur de contenu dans le processus plutôt que de compter sur un transfert passif de connaissances.

La formation doit également inclure une discussion sur les rôles et responsabilités du responsable et du animateur de contenu, sur la gestion de la dynamique de groupe (en particulier lorsque des observateurs de statut supérieur sont présents, pour éviter qu'ils n'influencent indûment les décisions des panélistes) et doit fournir des détails sur ce qu'il faut faire et ne pas faire en matière de facilitation de discussions pendant et après l'exécution de chacune des tâches, comme indiqué dans le **tableau 10** (les mêmes règles s'appliquent pour répondre aux questions des panélistes et faciliter l'évaluation des pratiques).

L'objectif principal de la formation sera de garantir que les animateurs de contenu connaissent parfaitement le GPF et l'évaluation, car ils devront aider les panélistes à interpréter les deux et à couvrir les trois tâches : alignement, mise en correspondance et analyse comparative. Les animateurs principaux et de contenu sont chargés de communiquer la politique reliant les procédures aux panélistes, tandis que les animateurs de contenu sont chargés de renforcer la formation globale avec les panélistes pendant le travail de groupe. Les deux animateurs doivent savoir comment répondre aux questions des panélistes et faciliter les discussions pertinentes.

Si le temps entre la formation et l'atelier est suffisant, il peut être utile d'entreprendre une répétition des sections pertinentes de l'atelier avec les animateurs de contenu, les animateurs principaux agissant en tant que panélistes, pour garantir la compréhension.

F. TEST TECHNIQUE

Dans les ateliers hybrides et à distance, il est essentiel d'effectuer un test technique auprès de tous les sites et participants avant l'atelier. Cela permettra de dépanner et de disposer de suffisamment de temps pour mettre en œuvre des solutions de secours lorsque des problèmes ne peuvent pas être résolus et garantir que le début de l'atelier principal ne soit pas perturbé.

Les activités suivantes doivent être entreprises pendant le test technique :

- **Connectivité**– tous les sites et tous les participants disposent-ils d'une connexion appropriée ?
- **Audio** – tous les lieux et tous les participants disposent-ils de microphones et de haut-parleurs adaptés ?

- **Plate-forme** – tous les participants sont-ils familiarisés avec les fonctionnalités de la plate-forme numérique (par exemple, couper le son, lever la main, utiliser les fonctions de chat, passer d'une salle de discussion de petit groupe à une autre, etc.) ?
- **Macros** – si durant l'atelier, il est prévu d'utiliser des formulaires numériques contenant des macros, celles-ci sont-elles accessibles à tous les panélistes ?
- **Application de messagerie** : tous les participants ont-ils téléchargé l'application de messagerie choisie et peuvent-ils accéder à la discussion de groupe ?

Tableau 10 : Objectif de la discussion, choses à faire et à ne pas faire par tâche

Tâche	Objectif des discussions	À faire	À ne pas faire
Tâche 1 : Évaluation et alignement du GPF (les panélistes travaillent de manière indépendante)	Pour vous assurer que les panélistes ont compris la tâche, découvrez les défis auxquels ils ont été confrontés et déterminez également s'il existe des items qui ne correspondent pas au GPF et qui, par conséquent, n'ont pas besoin d'être notés.	<ul style="list-style-type: none"> • Assurez-vous que tous les panélistes ont la possibilité de s'exprimer, de partager leurs évaluations et de poser des questions. • Assurez-vous que tous les panélistes prennent en compte chacune des étapes d'alignement et que leurs explications sur la façon dont ils ont sélectionné « aucune correspondance », « correspondance partielle » ou « correspondance complète » ont du sens et démontrent une compréhension des concepts. • Explorez les désaccords entre l'alignement des panélistes sur les énoncés de connaissances et/ou de compétences et leur adéquation en demandant aux panélistes des deux côtés d'expliquer volontairement pourquoi ils ont évalué la façon dont ils l'ont fait. 	<ul style="list-style-type: none"> • Informez un panéliste ou sous-entendez qu'un panéliste a mal aligné un item. • Informez un panéliste ou sous-entendez qu'un panéliste a sélectionné le mauvais niveau de correspondance. • Choisissez des panélistes individuels pour leur demander pourquoi ils ont aligné X items sur X énoncés de connaissances et/ou de compétences.
Tâche 2 : Faire correspondre les items d'évaluation avec les GPL et les GPD (les panélistes travaillent ensemble en groupes)	Pour vous assurer que les panélistes ont compris la tâche, découvrez les défis auxquels ils ont été confrontés, assurez-vous qu'ils ont pris en compte ce qui rend un item facile/difficile et assurez-vous également que le groupe est parvenu à un consensus sur le GPL et les GPD qui correspondent à chaque élément.	<ul style="list-style-type: none"> • Assurez-vous que tous les panélistes ont l'occasion de s'exprimer, de donner leur avis sur leur accord ou leur désaccord avec le consensus du groupe et de poser des questions. • Assurez-vous que tous les panélistes prennent en compte chacune des étapes de correspondance et que leurs explications sont claires et conformes à la méthodologie en ce qui concerne la façon dont ils ont sélectionné le GPL le plus bas auquel les apprenants devraient avoir les connaissances et/ou les compétences nécessaires pour répondre à un item. • Mettez en avant des points supplémentaires qui pourraient rendre un item facile ou difficile que les panélistes n'ont pas identifié. 	<ul style="list-style-type: none"> • Informez les panélistes ou sous-entendez que les panélistes ont mal associé un item à un GPL/GPD ou que leurs arguments sur ce qui rend un item facile/difficile sont faux.
Tâche 3, série 1 – Noter les items à l'aide de la méthode Angoff (les panélistes)	Pour vous assurer que les panélistes ont compris la tâche, demandez-leur d'expliquer pourquoi ils ont évalué un item	<ul style="list-style-type: none"> • Assurez-vous que tous les panélistes ont la possibilité de parler, de fournir des explications sur la façon dont ils ont évalué les items et la raison, et de poser des questions. • Assurez-vous que tous les panélistes 	<ul style="list-style-type: none"> • Informez les panélistes ou sous-entendez que les panélistes ont mal évalué un item. • Choisissez des panélistes individuels pour leur demander pourquoi ils ont évalué ainsi X items (Remarque : les

travaillent de manière indépendante)	de la manière dont ils l'ont fait. Leur explication doit faire référence au GPD et aux questions « serait » et « raisonnablement sûr ». Donnez également aux panélistes l'occasion de parler des désaccords sur les notes, car cela pourrait éclairer les décisions de notation de certains panélistes lors de la série 2.	prennent en compte chacune des étapes d'évaluation et que leurs explications sur les raisons pour lesquelles ils ont noté un item de la manière dont ils l'ont fait, font référence aux GPD, leur conceptualisation des apprenants à chacun des GPL, les points qui rendent l'item facile/difficile, et si ils sont « raisonnablement sûrs ». <ul style="list-style-type: none"> • Identifiez les items sur lesquels les panélistes ne sont pas d'accord et demandez aux panélistes volontaires qui ont noté « non » d'expliquer pourquoi et vice versa. • Encouragez les panélistes à considérer la difficulté de l'item et les données d'impact et à décider si cela affecte leurs jugements concernant la série 2. 	notes des panélistes sont censées être confidentielles, c'est pourquoi elles sont présentées au groupe par numéro de panéliste plutôt que par nom). <ul style="list-style-type: none"> • Cela implique que, parce que les données sur la difficulté des items montrent que les apprenants ont trouvé un item difficile, il devrait être noté « non ». Il est possible que de nombreux apprenants qui ont passé l'évaluation ne satisfaisaient tout simplement pas aux exigences des GPL.
Tâche 3, série 2 – Notation des items à l'aide de la méthode Angoff (les panélistes travaillent de manière indépendante)	Obtenez les réactions des panélistes par rapport à leurs références finales et aux données d'impact.	<ul style="list-style-type: none"> • Assurez-vous que tout le monde a la possibilité de s'exprimer et de poser des questions. 	<ul style="list-style-type: none"> • Formulez des affirmations non fondées sur la manière dont le gouvernement/l'agence d'évaluation régionale ou internationale utilisera les critères de référence.

CHAPITRE V

CHAPITRE V. MISE EN ŒUVRE DE L'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

Alors que le **chapitre III** fournit une explication sur la méthodologie utilisée dans l'atelier consacré à l'établissement de liens entre les politiques, ce chapitre fournit des orientations et des conseils aux animateurs sur la manière de diriger l'atelier et quand faire quoi.

Lorsque la langue parlée par les animateurs principaux n'est pas la langue maternelle des panélistes, il sera important de garantir la présence d'interprètes pour faciliter l'encadrement par une traduction simultanée. Ceci est important même lorsque les panélistes ont une compréhension de base du langage du animateur principal. Compte tenu de la nature technique des concepts impliqués dans l'établissement de liens entre les politiques et du niveau de discussion requis, il sera plus facile pour les panélistes d'exprimer leurs idées dans la langue de leur choix. Les animateurs principaux peuvent avoir besoin d'une formation sur la façon de dispenser l'atelier avec traduction simultanée si ce n'est pas quelque chose qu'ils ont déjà fait.

Comme décrit aux **chapitres III** et **IV**, les animateurs dirigeront des présentations et des activités au cours d'un certain nombre de séances. Pendant ce temps, ils présenteront la méthodologie de l'atelier, le GPF et l'évaluation lors de la familiarisation, puis dirigeront les panélistes à travers les trois principales tâches de l'établissement de liens entre les politiques :

- **Tâche 1** : vérifier l'alignement du contenu entre les évaluations et le GPF en utilisant une procédure standardisée
- **Tâche 2** : faire correspondre les items d'évaluation avec le GPF, c.-à-d. les GPL et les GPD
- **Tâche 3** : définir trois critères de référence mondiaux pour chaque évaluation en utilisant une méthode standardisée (une version modifiée de la procédure Angoff)²¹

Le modèle dans l'**annexe P** fournit les horaires de chaque activité pendant l'atelier. L'équipe du projet devra les organiser en jours/séances en fonction du type d'atelier (en présentiel, hybride ou à distance). Des exemples d'ordres du jour pour un atelier en personne et à distance se trouvent à l'**annexe Q**. Le **tableau II** présente les activités de l'atelier, liées aux 20 séries de diapositives de l'**annexe G** qui sont également organisées par activité et peuvent être utilisées pour créer des diapositives au jour le jour ou session par session en fonction de l'ordre du jour créé par l'équipe du projet. Bien que les horaires puissent varier légèrement pour répondre aux exigences locales, l'ordre des activités doit rester constant et les équipes de projet doivent s'assurer de disposer de suffisamment de temps pour permettre aux panélistes de bien comprendre ce que l'on attend d'eux et d'effectuer les tâches. L'équipe du projet devra également s'assurer qu'il y a suffisamment de temps dans l'ordre du jour pour que les animateurs puissent entreprendre les activités nécessaires entre les sessions concernées (par exemple, rassembler des données et produire des diapositives sur l'impact des données).

Les présentations sont animées en plénière par les animateurs principaux, et les activités sont dirigées en groupes (panels) par les animateurs de contenu. Les calculs de références et d'indicateurs doivent être effectués par les animateurs principaux et l'analyste de données. Les animateurs principaux et de contenu doivent organiser des discussions de contrôle ou administrer de courtes évaluations avec les panélistes à

²¹ Notez que si au cours des étapes 1, 2 ou 3, le gouvernement décide qu'il souhaite uniquement fixer un critère de référence pour le niveau « Atteint » ou si le gouvernement/l'agence d'évaluation ou le comité de l'examen de l'objectif 4.1.1 décide que l'évaluation est trop courte pour accueillir trois critères de référence aux trois principaux GPL, les panélistes n'ont alors qu'à définir un seul critère de référence (au lieu de trois) pour chaque évaluation.

la fin de chaque journée/session (plus de détails sont **inclus dans la** sous-section « Formulaire d'évaluation de l'atelier »). Indépendamment de ce qui est décidé pour les contrôles/évaluations réguliers, les panélistes doivent remplir une évaluation écrite de toutes les activités avant la fin de l'atelier à des fins de rapport.

Les animateurs doivent se réunir à la fin de la journée/session pour préparer la journée/session suivante au cas où des changements seraient nécessaires dans les plans à la suite des leçons apprises.

Tableau 11 : Résumé des tâches et des activités de l'atelier consacré à l'établissement de liens entre les politiques

Tâche	Présentation	Activité
Ouverture	1	Mots de bienvenue et présentations
		Allocution des représentants du ou des gouvernements, de l'agence d'évaluation (le cas échéant) et de l'organisme donateur (le cas échéant)
Tâche	Présentation	Activité
Familiarisation	2	Présentation du programme, des objectifs et aperçu général de la méthodologie
	3	Familiarisation avec le GPF
Tâche 1 – Alignement	4	Familiarisation avec l'outil d'évaluation
	5	Former les panélistes à l'exercice d'alignement (y compris les exercices pratiques)
	6	Les panélistes entreprennent une activité d'alignement (activité indépendante)
Tâche 2 – Mise en correspondance	7	Présentation et discussion sur les résultats de l'alignement
	8	Former les panélistes à la tâche de mise en correspondance
	9	Les panélistes entreprennent une activité de mise en correspondance (activité de groupe)
Tâche 3 – Analyse comparative	10	Présentation et discussion des résultats de la mise en correspondance
	11	Présentation des normes mondiales et de l'approche d'analyse comparative
	12	Former les panélistes à la méthode Angoff
	13	Les panélistes entreprennent la méthode Angoff avec des exercices pratiques
	14	Série 1
	15	Présentation et discussion des résultats de la série 1 et des données d'impact
	16	Présentation de la série 2 Angoff
17	Série 2	
Évaluation	18	Présentation des résultats de la série 2
Clôture	19	Évaluation de l'atelier
Documentation	20	Mots de clôture et présentation des certificats
	Après l'atelier	Réalisation de la documentation technique

Des informations sur chacune des présentations ci-dessus (1 à 20) sont fournies ci-dessous, ainsi que des conseils à l'intention des animateurs.

A. OUVERTURE

1. Mots de bienvenue et présentations

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 1), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (VOIR CHAPITRE III, SECTION D)

Dans cette présentation, vous vous présenterez et formulerez vos remarques liminaires. Vous devrez inviter les représentants du gouvernement et tous les responsables de l'éducation des donateurs, le cas

échéant, à faire des remarques liminaires. Le partenaire de mise en œuvre peut également faire des remarques si un projet co-parraine l'atelier. Les participants à l'atelier et l'équipe du projet se présenteront. Vous identifierez le matériel d'atelier trouvé dans les dossiers d'atelier des panélistes. Vous discuterez de la logistique de l'atelier, qui variera en fonction du type d'atelier (en présentiel, à distance ou hybride), mais pourra inclure des informations relatives au lieu, aux salles plénières et en petits groupes, à l'hébergement, aux repas, aux indemnités journalières, au transport, à la technologie et aux méthodes de communication.

Figure 17 : Conseils aux animateurs pour la présentation d'ouverture

Les responsables gouvernementaux/personnels des agences d'évaluation, les fonctionnaires donateurs de l'éducation et les partenaires de mise en œuvre devraient disposer d'environ 10 minutes chacun pour soumettre leurs remarques. Au fur et à mesure que chaque panéliste se présente au groupe, vous pouvez leur demander d'indiquer leur nom, leur emplacement/lieu de travail et leur emploi. Après la présentation générale, prévoyez un temps d'environ 10 minutes pour les questions-réponses. Assurez les participants que les présentations formelles ne sont qu'un aperçu et que les sessions suivantes approfondiront chacun des sujets mentionnés.

2. Présentation du programme, des objectifs et aperçu général de la méthodologie

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 2), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES

Dans cette présentation, vous fournirez aux participants un aperçu du programme de l'atelier, des informations générales sur la méthode d'établissement de liens entre les politiques, les indicateurs de l'ODD 4.1.1, les indicateurs « F » de l'USAID (le cas échéant) et le GPF. Vous expliquerez brièvement la nécessité de critères de référence qui détermineront la compétence minimale mondiale lors des évaluations. Vous expliquerez les trois tâches d'établissement de liens entre les politiques : 1) vérifier l'alignement, 2) faire correspondre les items d'évaluation avec les niveaux de compétence et les descripteurs, et 3) définir les références mondiales à l'aide d'une méthode standardisée.

Figure 18 : Conseils aux animateurs sur la présentation du contexte

Lors de la présentation du GPF et du PLT, fournissez le contexte de l'atelier en donnant un bref historique et en décrivant les activités futures. Utilisez le graphique avec l'échelle GPF, comprenant les quatre niveaux de compétence et les trois points de référence. Expliquez que l'objectif de l'atelier est de fixer des critères. Les critères de référence seront utilisés pour comparer les résultats des évaluations entre les pays, agréger les résultats des évaluations pour les rapports mondiaux et suivre les progrès au fil du temps. Dites aux panélistes que plus d'informations seront fournies au cours de chaque session.

B. FAMILIARISATION

Lorsque cela est possible, il est idéal d'entreprendre les exercices de familiarisation avant l'atelier pour donner aux panélistes le temps d'intérioriser l'information avant de l'utiliser au cours de l'atelier. Cependant, cela peut ne pas être possible, en particulier pour les ateliers réalisés en présentiel.

3. Familiarisation avec le GPF

MATÉRIEL : DIAPOSITIVES D'ANIMATION (PRÉSENTATION 3), EXTRAITS DE NIVEAU/MATIÈRE PERTINENTS DU GPF

Dans cette présentation, vous présenterez le GPF, notamment en présentant chacun des domaines, concepts, sous-concepts, énoncés de connaissances et/ou de compétences, ainsi que les GPL et GPD. Vous fournirez des informations générales sur le développement du GPF et parcourez tous les GPD pour le niveau scolaire concerné (2^{ème} année pour l'ODD 4.1.1 (a), 5^{ème} année pour l'ODD 4.1.1 (b) et 8^{ème} année pour l'ODD 4.1.1(c)). Vous discuterez de termes prêtant à confusion et demanderez aux panélistes de donner des exemples d'items qui pourraient être utilisés pour mesurer la norme de performance décrite dans le GPD. Lorsque la note de l'évaluation diffère de celle utilisée pour l'établissement de liens entre les politiques, vous devez également fournir la note de l'évaluation issue du GPF, car elle sera probablement nécessaire ultérieurement pour l'exercice de mise en correspondance.

Figure 19 : Conseils aux animateurs sur la présentation du GPF

Lorsque la familiarisation a lieu dans l'atelier, divisez les informations en fonction des besoins (par exemple, jusqu'à et y compris les énoncés de connaissances ou de compétences avant la tâche d'alignement, et les GPD et GPL entre la tâche d'alignement et celle de mise en correspondance) pour éviter la surcharge dès le premier jour.

Assurez-vous de consacrer suffisamment de temps à examiner chacun des termes clés et les GPD pour garantir la compréhension des panélistes. Vous souhaitez peut-être demander aux animateurs de contenu de traduire certains termes dans la langue locale pour garantir que tout le monde comprenne la même chose. Prenez également le temps de faire une pause lors de l'examen de chaque GPD pour engager les panélistes dans une discussion sur ce GPD et sur les types d'items d'évaluation qu'ils pourraient envisager d'utiliser pour le mesurer. S'assurer qu'il est clair que lorsque vous parlez d'atteindre les compétences minimales mondiales lors de l'atelier, vous parlez d'apprenants qui possèdent les compétences définies dans le GPF.

Il est important de veiller à ce que cette session soit engageante pour les panélistes, et qu'ils ne se contentent pas seulement d'écouter les animateurs. La Communauté de pratique pour l'établissement de liens entre les politiques a partagé du matériel, y compris des activités pour les panélistes, pour faciliter la familiarisation, ce qui peut être utile.

4. Familiarisation avec l'outil d'évaluation

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 4), OUTIL D'ÉVALUATION

(Remarque : vous devrez créer des diapositives supplémentaires pour cette présentation ; nous vous conseillons d'élaborer une diapositive par item d'évaluation, par paire ou par item)

Dans cette présentation, vous présenterez l'outil d'évaluation, décrierez comment il est administré, comment il est noté et à quoi ressemblait l'échantillon de population pour la dernière itération de l'évaluation (p. ex., de quelle zone/population était-il représentatif). Vous passerez en revue chacun des items de l'évaluation et vous assurerez que les panélistes les comprennent. Vous pouvez également demander aux panélistes de faire passer l'évaluation les uns aux autres (pour les évaluations administrées individuellement) ou de passer l'évaluation eux-mêmes (pour les évaluations administrées en groupe) pour garantir une meilleure compréhension.

Figure 20 : Conseils aux animateurs sur la présentation de l'évaluation

Cette présentation doit être dirigée par une personne connaissant parfaitement l'évaluation. La séance doit être interactive et ne pas se contenter de montrer l'évaluation aux panélistes.

Assurez-vous de consacrer suffisamment de temps à chaque item d'évaluation pour vous assurer que les panélistes les comprennent, la manière dont ils sont administrés et quelles pourraient être certains obstacles courants. Lors de l'examen de l'activité préalable à l'atelier, assurez-vous que les panélistes ont sélectionné les apprenants à évaluer en fonction de ceux dont ils savaient qu'ils possédaient les connaissances et/ou les compétences décrites dans le GPF pour un niveau et un GPF particuliers. Si tel est le cas, les scores de ces apprenants peuvent s'avérer particulièrement utiles aux panélistes pour établir des critères. Si les panélistes n'ont pas été en mesure d'évaluer les apprenants qui répondent aux définitions du GPF pour atteindre partiellement, atteindre ou dépasser la compétence minimale mondiale, les scores des apprenants qu'ils ont évalués sont moins importants et ils devraient plutôt simplement utiliser les résultats de cette activité pour éclairer leur compréhension de la difficulté des items et des procédures d'administration des tests. Prenez suffisamment le temps de poser des questions et de discuter de l'évaluation.

C. TÂCHE 1 - ALIGNEMENT

5. Former les panélistes à l'exercice d'alignement (y compris les exercices pratiques)

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 5), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE H)

Dans cette présentation, vous revisiterez le GPF, plus particulièrement les sous-concepts et les énoncés de connaissances et/ou de compétences. Vous décrivez le processus en trois étapes dans lequel les panélistes s'engageront pour vérifier l'alignement des évaluations avec les énoncés de connaissances et/ou de compétences décrits par le GPF (voir la section sur la **tâche I** du chapitre III et le tableau 3 du GPF) et le processus que les animateurs utiliseront pour résumer les résultats. Vous expliquerez les trois niveaux d'alignement ou de correspondance – correspondance complète, partielle et sans correspondance

– en comptant à la fois complet et partiel pour l’alignement. Vous pouvez expliquer la méthode standardisée pour déterminer le niveau d’étendue et de profondeur d’alignement entre la ou les évaluations et le GPF. Vous présenterez aux participants quelques exemples d’éléments pour vous assurer qu’ils comprennent la tâche. Des exemples d’items de lecture et de mathématiques sont inclus dans la **présentation 5** que vous pouvez utiliser à cette fin, ou vous pouvez sélectionner/développer les vôtres. Notez que les exemples d’items ne doivent pas être trop similaires aux items d’évaluation réels que les panélistes évalueront, car cela pourrait biaiser les notes, mais il est utile qu’ils couvrent des sous-concepts similaires. Enfin, vous pouvez partager les critères d’alignement répertoriés dans les **tableaux 4 et 5**, bien que cela ne soit pas obligatoire car l’alignement global rapporté sera tiré de l’exercice effectué dans le cadre de l’auto-évaluation.

Figure 21 : Conseils aux animateurs sur la présentation de l’alignement

Lorsque vous décrivez l’activité d’alignement, rappelez aux panélistes que le GPF a été développé comme un ensemble global de connaissances et de compétences et de GPD associés, tirés d’un contenu mondial consensuel. Assurez-vous que les panélistes connaissent la différence entre les énoncés de connaissances et/ou de compétences et les GPD (normes de contenu et de performance). Parcourez attentivement les exemples et chacune des deux étapes et sous-étapes décrites dans la section sur la **tâche 1 du chapitre III**. Dites aux panélistes que certains items d’évaluation peuvent ne pas correspondre au GPF puisque chaque pays a ses propres normes. C’est tout à fait normal. Assurez-vous qu’ils comprennent que les deux items avec une correspondance partielle ou complète sont pris en compte dans les critères d’alignement. Lorsque la note de l’évaluation est différente de la note du GPF utilisée pour l’établissement de liens entre les politiques, assurez-vous d’expliquer pourquoi c’est le cas, pour garantir que la même norme soit appliquée dans tous les pays.

6. Les panélistes entreprennent une activité d’alignement (activité indépendante)

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 6), DOSSIERS D’ATELIER À L’INTENTION DES PANÉLISTES (EN PARTICULIER L’ANNEXE H)

Dans cette activité, vous donnerez aux panélistes l’occasion de poser des questions avant de procéder à l’alignement des items d’évaluation avec les énoncés de connaissances et/ou de compétences du GPF à l’aide de la feuille d’inscription fournie.

Figure 22 : Conseils à l’intention des animateurs pour la tâche 1 – Aligner l’évaluation ou les évaluations sur le GPF

Bien que la discussion soit encouragée pendant le travail de groupe, chaque panéliste doit procéder à ses propres évaluations d’alignement individuelles et indépendantes, ou à des évaluations d’items de connaissances et/ou de compétences, et soumettre son formulaire aux animateurs de contenu pour analyse par les animateurs principaux ou les analystes de données. Les panélistes doivent uniquement s’aligner sur les énoncés de connaissances et/ou de compétences pertinents pour le niveau scolaire (2^{ème} année pour l’ODD 4.1.1(a), 5^{ème} année pour l’ODD 4.1.1(b) et 8^{ème} année pour l’ODD 4.1.1(c), comme illustré par chaque « x » dans le tableau 3 du GPF.

7. Présentation et discussion des résultats de l’alignement

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 7)

(Il est recommandé de créer des diapositives supplémentaires pour cette présentation, y compris une diapositive par item pour lequel un désaccord important est survenu entre les panélistes sur les énoncés de connaissances et/ou de compétences avec lesquels l’item correspond).

Dans cette présentation, vous couvrirez les résultats de l’activité d’alignement. Vous pouvez également aborder le niveau d’alignement atteint en fonction des critères d’alignement présentés dans les **tableaux 4 et 5**. Vous souhaitez également examiner les items individuels et les notes d’alignement pour lesquels un désaccord considérable est survenu entre les panélistes sur les énoncés de connaissances et/ou de

compétences auxquels l’item correspondait. Des conseils pour faciliter cette discussion sont inclus dans le **tableau 10** ci-dessus dans la section Formation des animateurs de contenu.

Figure 23 : Conseils à l’intention des animateurs pour l’examen des résultats de la tâche I

Bien qu’un accord ne soit pas requis sur la tâche d’alignement, il l’est pour la mise en correspondance. Les animateurs de contenu doivent avoir accès aux résultats d’alignement convenus issus du processus d’auto-évaluation pour aider à orienter les discussions.

D. TÂCHE 2 : MISE EN CORRESPONDANCE

8. Former les panélistes à la tâche de mise en correspondance

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 8), DOSSIERS D’ATELIER À L’INTENTION DES PANÉLISTES

Dans cette présentation, vous vous baserez sur l’alignement réalisé au cours de la tâche I (avec les énoncés de connaissances et/ou de compétences) pour discuter de la mise en correspondance avec les GPL et aux GPD (également appelées normes de performance) – les points de vue sur l’alignement peuvent changer au cours de l’exercice de mise en correspondance, qui est autorisé à condition qu’un consensus soit atteint à la fin. Vous guiderez les panélistes en répondant aux trois questions requises dans le cadre de la tâche (voir la section sur la **tâche 2** du chapitre III pour les questions) - à savoir, quelles connaissances et/ou compétences sont requises pour répondre correctement à l’item, qu’est-ce qui rend l’item facile / difficile, et quel est le GPL le plus bas correspondant à l’item. Pour la lecture, cela inclura une discussion sur l’annexe sur la complexité des passages en lecture du GPF, car cela devra également être pris en compte dans le cadre de la tâche de mise en correspondance, pour s’assurer que le texte est adapté au niveau scolaire. Vous présenterez aux participants quelques exemples d’items pour vous assurer qu’ils comprennent la tâche. Il existe des exemples d’items inclus dans la **présentation 8** pour la lecture et les mathématiques que vous pouvez utiliser à cette fin, ou vous pouvez sélectionner ou développer les vôtres.

Vous souhaitez peut-être également que les panélistes aient accès à leurs propres documents de programme pour leur rappeler les attentes au niveau scolaire. C’est particulièrement le cas en lecture, où le GPF contient des références, par exemple, à des mots courants pour le niveau. Cependant, si vous partagez des documents pédagogiques, assurez-vous que tous les panélistes indiquent clairement qu’ils relient les éléments au GPF et non à leur propre programme, dont le contenu peut être lié à différentes tranches d’âge.

Figure 24 : Conseils à l’intention des animateurs sur la présentation de la mise en correspondance de la tâche 2

Rappelez aux panélistes que cette activité s’appuie sur la compréhension des items d’évaluation et du GPF acquis grâce à l’activité d’alignement. Le concept clé est de faire correspondre les items avec les GPL et GPD les plus bas qui décrivent les attentes que les apprenants doivent satisfaire pour répondre correctement à l’item pour le niveau considéré. Si le groupe a évalué l’item comme correspondant partiellement, il devra prendre en compte les deux GPD pertinents et probablement sélectionner le plus élevé des deux GPL, car les apprenants doivent répondre aux attentes des deux pour répondre correctement à l’item. Si le groupe a évalué l’item comme étant « sans correspondance », les panélistes doivent suivre le processus décrit dans la **Figure 10**.

Si un item correspond à un descripteur d’un niveau autre que celui considéré (2^{ème} année pour l’ODD 4.1.1(a), 5^{ème} année pour l’ODD 4.1.1(b) et 8^{ème} année pour l’ODD 4.1.1(c), cela doit également être enregistré, en indiquant la note correspondante. Ces informations seront utilisées lors de la tâche d’analyse comparative. Si l’item est lié à un niveau inférieur, il est probable que les apprenants du niveau considéré soient en mesure de répondre correctement. Si l’item est lié à un niveau supérieur, il est peu probable que les apprenants du niveau considéré soient en mesure d’y répondre correctement.

9. Les panélistes entreprennent une activité de mise en correspondance (activité de groupe)

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 9), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES

Dans cette activité, vous opérationnaliserez la présentation. Vous offrirez aux panélistes l'occasion de poser des questions sur les GPL et les GPD. Vous clarifierez à nouveau la différence entre les énoncés de connaissances et/ou de compétences et les GPD. Vous diviserez le panel en groupes distincts au niveau du panel pour chaque évaluation (niveau, matière et langue) liées tout au long de l'atelier, et les animateurs de contenu les guideront en faisant correspondre chaque élément avec les GPL et GPD les plus bas. Les animateurs de contenu travailleront également pour les aider à parvenir à un consensus. Lorsque cela s'avère difficile, vous devez utiliser le résultat de l'activité d'alignement menée dans le cadre de l'auto-évaluation pour guider la discussion, car en fin de compte, il a été déterminé que c'était l'alignement correct par le pays.

Figure 25 : Conseils à l'intention des animateurs pour superviser l'activité de mise en correspondance de la tâche 2

Assurez-vous que les panélistes examinent item par item et discutent de la manière dont les items correspondent aux GPD les plus bas. Il peut être utile que les panélistes discutent de leurs correspondances en petits groupes, puis se réunissent pour parvenir à un consensus au sein de leurs panels. Rappelez-leur d'écrire les réponses aux trois questions de la tâche directement sur leur outil d'évaluation/livret de test à côté de l'item.

10. Présentation et discussion des résultats de la mise en correspondance (Tâche 2)

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 10), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES

Dans cette présentation, vous fournirez les résultats de la mise en correspondance et vérifierez la compréhension des panélistes du processus de mise en correspondance. Vous résumerez les réponses consensuelles aux trois questions de cette activité. Étant donné que le processus de mise en correspondance est une activité de groupe, vous n'aurez peut-être pas besoin de consacrer beaucoup de temps à examiner les résultats. Vous pourriez simplement demander si les panélistes se sont concentrés sur les GPD pour prendre leurs décisions, s'il y a eu des désaccords, et si oui, comment ils ont été résolus. Un exemple où vous voudriez consacrer beaucoup de temps à cette activité est si vous avez deux panels différents fixant des critères de référence sur une seule évaluation, probablement à des niveaux scolaires différents. Si tel est le cas, l'alignement vertical entre les critères de référence sera essentiel, et l'examen des correspondances GPD pourrait aider à indiquer les défis qui peuvent survenir dès le début (par exemple, si un panel de troisième année associe un item à un niveau scolaire inférieur à celui du panel de deuxième année). Des conseils supplémentaires pour faciliter cette discussion sont inclus dans le **tableau 10**.

Figure 26 : Conseils à l'intention des animateurs pour l'examen des résultats de la mise en correspondance de la tâche 2

Les panélistes devront se mettre d'accord sur les correspondances, c.-à-d. parvenir à un consensus, avant de passer au processus d'analyse comparative. Notez que les tâches 1 et 3 impliquent des évaluations individuelles et indépendantes, mais que la tâche 2 implique un consensus entre les panélistes sur les mises en correspondance. Assurez-vous que les résultats des mises en correspondance sont enregistrés par chaque panéliste dans son outil d'évaluation/livret de test.

E. TÂCHE 3 - ANALYSE COMPARATIVE

11. Présentation des normes mondiales et de l'approche d'analyse comparative

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 11), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES

Dans cette présentation, vous expliquerez les principaux concepts derrière l'analyse comparative mondiale en relation avec le GPF à l'aide de plusieurs exemples. Vous expliquerez le premier graphique (diapositive 84) montrant le critère de conformité sur les deux échelles – évaluation nationale et GPF – et comment les critères relient les échelles aux points de score identifiés. Vous expliquerez le graphique qui montre trois évaluations nationales avec des critères différents en fonction de la difficulté de ces évaluations (diapositive 85). Vous couvrirez le troisième graphique de la présentation (diapositive 86) avec les pourcentages d'apprenants dans les GPL (catégories) à partir des ensembles de données d'évaluation, qui sont utilisés pour les comparaisons, l'agrégation et le suivi des indicateurs de l'ODD 4.1.1 et de l'USAID.

Figure 27 : Conseils aux animateurs pour la présentation de l'analyse comparative mondiale

Cette présentation procède étape par étape à travers les échelles d'évaluation et le graphique GPF, avec d'un critère de référence (deux niveaux et pourcentages) à trois critères (quatre niveaux et pourcentages). Assurez-vous que les panélistes réalisent que l'emplacement des critères dépend de la difficulté de l'évaluation. Ils doivent également savoir que chaque évaluation a un niveau de difficulté différent et a donc des critères différents par rapport à l'échelle commune.

12. Former les panélistes à la méthode Angoff

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 12), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE H)

Dans cette présentation, vous expliquerez le processus standardisé d'établissement de critères à l'aide de la version Oui-Non de la méthode Angoff (voir la section sur la **tâche 3**). Vous fournirez des informations générales sur la méthode Angoff et sur la manière dont elle est utilisée pour établir des critères mondiaux sur les évaluations nationales et internationales. Vous présenterez l'idée de deux séries d'évaluations d'items. Vous indiquerez que les panélistes doivent procéder à des évaluations individuelles et indépendantes de chaque item pour définir leurs critères de référence, qui sont ensuite moyennés pour calculer les critères du panel. Vous montrerez aux panélistes comment leurs critères individuels sont calculés, en additionnant le nombre de décisions « Oui » pour chaque GPL (en notant que si, par exemple, une décision « Oui » est enregistrée pour le GPL « atteint », alors elle sera également enregistré pour le GPL « dépasse »). Vous expliquerez également que le critère global du panel sera calculé en prenant la moyenne de tous les critères individuels.

Si l'évaluation comporte des items polytomiques, assurez-vous d'inclure les diapositives pertinentes de la **présentation 12** pour expliquer qu'elles doivent déterminer si les apprenants de chaque GPL obtiendront chaque point à tour de rôle.

Figure 28 : Conseils aux animateurs pour la présentation de la méthode Angoff de la tâche 3

Indiquez aux panélistes que le même processus se produit pour les critères initiaux (Série 1) et les critères finaux (Série 2). Présentez les concepts d'attentes de l'apprenant (« devrait » selon les GPD et des attentes réalistes, et « serait », basé sur la réalité dans des situations de test) ainsi que la nécessité de fixer les critères de référence au GPL le plus bas qui correspond aux attentes que les apprenants doivent satisfaire pour répondre correctement à l'item. Un organigramme des évaluations et des exemples est fourni aux panélistes dans les diapositives et dans la **Figure 8**, ainsi que des conseils d'évaluation.

Fournissez une description claire des apprenants aux niveaux suivants : juste partiellement atteint (JPA), juste atteint (JA) et juste dépasse (JD) par rapport au GPF et comment cela se rapporte à leurs propres apprenants. Précisez que les élèves de leur classe peuvent ne pas être représentatifs de ceux décrits dans le GPF, en raison de choix différents effectués dans le programme ou de circonstances spécifiques au pays, par exemple.

Si certains éléments ont été classés comme « sans correspondance », la discussion dans l'exercice de correspondance (voir **Figure 10**) devrait aider les panélistes à déterminer si les apprenants répondraient correctement à l'item, en fonction de leur expérience.

13. Les panélistes adoptent la méthode Angoff sur les exercices de pratique

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 13), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE I)

Dans cette activité, vous reverrez les présentations sur l'analyse comparative mondiale et la méthode Angoff dans les panels. Vous passerez en revue les exemples de la présentation et du logigramme, avec les notations Angoff. Vous accorderez suffisamment de temps aux panélistes pour mettre en pratique leurs évaluations des items à l'aide d'échantillons d'items présélectionnés. Des exemples d'items de lecture et de mathématiques sont inclus dans la **présentation 13** que vous pouvez utiliser à cette fin, ou vous pouvez sélectionner/développer les vôtres. Notez que les exemples d'items ne doivent pas être trop similaires aux items d'évaluation réels que les panélistes évalueront, car cela pourrait biaiser les notes, mais il est utile qu'ils couvrent des sous-concepts similaires. Vous dirigerez les discussions sur les notes des panélistes au sein du panel. Vous donnerez aux panélistes l'occasion de poser des questions et de clarifier le processus.

Figure 29 : Conseils aux animateurs sur la pratique Angoff de la tâche 3

Insistez sur le fait qu'une partie clé de cette activité repose sur la mise en correspondance de la tâche 2, dans laquelle les panélistes ont fait correspondre leurs items avec les GPL et GPD les plus bas du GPF. Ces correspondances fournissent des informations permettant d'évaluer les exemples d'items (en supposant que les mêmes exemples d'items ont été utilisés tout au long) et, plus important encore, les items réels de l'activité suivante. Ils doivent s'assurer qu'ils correspondent à la fois aux énoncés de connaissances et/ou de compétences (tâche 1) et aux GPD (tâche 2), et considérer ce qui rend un item facile ou difficile (de la tâche 2) et s'ils sont raisonnablement sûrs qu'un apprenant peu compétent y répondrait correctement. Les panélistes doivent être clairs sur le processus d'évaluation des items avant de passer à la série 1. Vous devez prévoir suffisamment de temps pour les questions au cours de cette session.

14. Série 1

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 14), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE I)

Dans cette activité, vous guiderez les panélistes dans l'application de la méthode Angoff pour noter les items d'évaluation. Vous expliquerez le formulaire d'évaluation des items (comme indiqué dans le **tableau 7**) qu'ils remplissent pour les séries 1 et 2. Vous réitérerez que les panélistes doivent évaluer les items individuellement et indépendamment, ce qui est différent de l'activité de correspondance dans laquelle ils sont parvenus à un consensus. Vous expliquerez aux panélistes qu'un écart entre eux est attendu, mais il doit être basé sur une compréhension commune des éléments et du GPF – lié aux résultats de la tâche de mise en correspondance, à laquelle tous les panélistes devraient avoir accès. Vous

montrerez aux panélistes comment calculer leurs propres critères de référence. Puis, à la fin de la journée, l'analyste de données vérifiera ces calculs et fera la moyenne entre les panélistes afin de générer des références pour les panels (voir **l'annexe T** pour plus de détails sur ces calculs). Les panélistes réaliseront leurs évaluations de la série 1 individuellement, mais pourront poser des questions individuelles aux animateurs pendant le processus.

Figure 30 : Conseils aux animateurs pour superviser la tâche 3 – Notations de la série 1

Les panélistes doivent savoir qu'ils doivent prendre leur temps quand à la notation de la série 1. Ils doivent être pleinement conscients que la collaboration avec les autres panélistes n'est pas acceptée dans cette activité, mais qu'ils auront l'occasion de discuter de leurs notes avec d'autres panélistes avant la dernière série (2e série). Les panélistes doivent s'assurer qu'ils correspondent aux énoncés de connaissances et/ou de compétences du GPF et des GPD. Il est également important qu'en répondant aux questions des panélistes, les animateurs se contentent de fournir des conseils sur la méthodologie, mais pas d'orienter les panélistes sur la manière d'évaluer un item particulier.

15. Présentation et discussion des résultats de la série 1 et des données d'impact

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 15), DOSSIERS D'ATELIER

ANNEXES N, ANNEXES O ET ANNEXES T

Dans cette présentation, vous expliquerez en détail les analyses des critères de la série 1 (tous présentés de manière anonyme, à l'aide des identifiants des panélistes) : 1) les critères des panélistes individuels et leurs distributions, 2) les critères au niveau du panel (voir **l'annexe T** pour plus de détails sur la façon de calculer les repères) et les informations normatives (statistiques de localisation) des repères des panélistes (les détails sur la façon de créer ce graphique sont inclus à **l'annexe O**), 3) les évaluations des items par rapport à leur difficulté réelle (voir **l'annexe N**), 4) moyennes des critères des panélistes, et 5) données d'impact avec pourcentages d'apprenants par GPL sur la base des références fixées par les panélistes lors de la première série. Vous engagerez les panélistes dans des discussions basées sur chacune de ces analyses. Voir le **tableau 10** pour obtenir des conseils sur la manière de mener cette discussion.

Figure 31 : Conseils aux animateurs pour partager les résultats de la série 1

Les analyses des diapositives génériques devront être remplacées par des analyses réelles basées sur les notes des panélistes lors de l'atelier. Discutez des différences dans les notes des panélistes et des raisons qui sous-tendent ces différences. Examinez les critères les plus élevés et les plus bas des panélistes. Vous souhaitez peut-être également examiner des items individuels pour lesquels un désaccord important est survenu. Demandez aux volontaires qui ont noté un item d'une manière d'expliquer pourquoi et aux volontaires qui l'ont noté d'une autre manière d'en donner la raison. L'idée est d'aider les panélistes à mieux comprendre les différentes options de notation afin de mieux éclairer leurs notes de la série 2. Des conseils pour cette discussion sont inclus dans le **tableau 11**. Demandez également aux panélistes de comparer les valeurs p réelles (statistiques de difficulté) avec leurs notes pour voir si leurs notes sont cohérentes avec les données – tout en n'oubliant pas de leur rappeler d'être prudents dans l'interprétation des valeurs p des items, car elles pourraient pas nécessairement être révélatrices d'une difficulté « mondiale ». Et enfin, demandez-leur si les données d'impact correspondent à ce qu'ils attendraient de la population évaluée. Découvrez pourquoi les résultats peuvent être différents de leurs attentes. Renforcez l'idée qu'ils doivent avoir une compréhension commune mais pas des notes communes, c'est-à-dire que la variation est normale et que les résultats sont moyennés pour calculer les critères de référence du panel.

S'il existe des panélistes atypiques dont les points de vue sont très différents de ceux de leurs pairs, vous devez avoir une conversation séparée avec eux pour vérifier qu'ils comprennent la tâche. Les valeurs aberrantes seront supprimées dans le cadre du processus final d'établissement des critères de référence.

16. Présentation de la série 2 Angoff

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 16), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE I)

Dans cette présentation, vous passerez brièvement en revue les procédures utilisées dans les notations de la série 1 comme guide pour la série 2. Vous expliquerez que les panélistes doivent examiner les notations de la série 1, prendre en compte les données et les discussions, puis réviser leurs notes pour la série 2 (ce n'est pas grave si les panélistes ne modifient pas leurs notes, mais ils doivent passer par le processus de révision de chaque item). Vous indiquerez aux panélistes qu'ils doivent utiliser la série 1 comme point de départ pour établir leurs évaluations de la série 2.

Figure 32 : Conseils aux animateurs pour la présentation de la série 2 Angoff

Tout changement dans les notes des panélistes entre la série 1 et la série 2 doit être basé sur un niveau accru de compréhension, tant pour les panélistes eux-mêmes que pour les panels. Cela devrait conduire les panélistes à devenir autonomes et à devenir des participants de groupe, avec l'idée qu'une meilleure compréhension devrait conduire à une plus grande précision et cohérence dans les critères de référence.

17. Série 2

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 17), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE I)

Dans cette activité, vous demanderez aux panélistes s'ils ont des questions sur la série 1 ou sur la présentation des résultats de la série 1. Vous demanderez aux panélistes de 1) rester concentrés sur le contenu de l'item par rapport aux GPL et GPD, 2) continuer à prendre en compte la difficulté des items comme base pour formuler leurs jugements, 3) apporter des ajustements, le cas échéant, à leurs notes de la série 1 en fonction de leurs jugements individuels et indépendants, et 4) leur rappeler de réfléchir à la manière dont les apprenants « répondraient » plutôt aux items que la façon dont ils « devraient » y répondre, et de s'assurer qu'ils sont au moins « raisonnablement sûrs » de leur note. Vous demanderez aux panélistes de soumettre leurs formulaires d'évaluation – les mêmes formulaires d'évaluation qu'à la série 1 – aux animateurs de contenu après avoir évalué les items de la série 2.

Figure 33 : Conseils aux animateurs pour la présentation de la série 2 Angoff

Il est important de surveiller les panélistes pendant qu'ils effectuent leurs évaluations de la série 2. Certains panélistes peuvent ne pas prendre correctement en compte les discussions et les données de la série 1. Ils doivent prendre leur temps et se rendre compte qu'il s'agit de leur dernière opportunité de procéder aux évaluations les plus précises possibles sur la base de leur connaissance des évaluations, du GPF, des données et des discussions.

18. Présentation des résultats de la série 2

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 18), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES ANNEXE T

Dans cette présentation, vous résumerez le processus qui a été suivi au cours de l'atelier et fournirez les critères finaux aux panélistes, avec des commentaires sur les changements entre la série 1 et la série 2. Vous fournirez les analyses suivantes : 1) Moyennes des critères des séries 1 et 2 des panélistes, c'est-à-dire les critères de référence du ou des panel(s), 2) une explication des changements entre les séries, et 3) les données d'impact sur le pourcentage d'apprenants dans les GPL. Vous présenterez les résultats sous forme de tableaux et de graphiques. Vous dirigerez une brève discussion sur les résultats en tant qu'activité technique finale de l'atelier, même si les panélistes ne pourront pas modifier leurs décisions à la suite d'une discussion.

Au cours de l'atelier, une analyse est effectuée à partir des résultats de tous les panélistes. Cependant, il peut encore y avoir des panélistes atypiques après la série 2. Le processus pour traiter ces cas est discuté au **chapitre VI**, mais cela devrait avoir lieu une fois l'atelier terminé afin de ne pas attirer l'attention sur les individus concernés pendant l'atelier.

Figure 34 : Conseils aux animateurs pour la présentation des résultats finaux

Les résultats sont plus limités que la présentation après la série 1. Le point principal est de comparer les changements entre la série 1 et la série 2, ainsi que de discuter si les panélistes estiment que les résultats sont raisonnables. Encore une fois, les animateurs principaux et l'analyste de données devront remplacer le tableau dans les diapositives en fonction des résultats de l'atelier.

F. ÉVALUATION

19. Activité – Évaluation de l'atelier

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 19), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES (EN PARTICULIER L'ANNEXE R)

Idéalement, l'évaluation aura été menée à la fin de chaque tâche pour permettre d'identifier et de résoudre les problèmes. Si cela n'est pas possible, une seule évaluation pourra être réalisée à la fin de l'atelier. Vous fournirez des instructions aux panélistes pour remplir le formulaire d'évaluation de l'atelier en fonction du moment où il est administré.

Les identifiants des panélistes seront collectés au cas où un panéliste indiquerait sur le formulaire d'évaluation qu'il n'est pas sûr de ses notes, ce qui pourrait remettre en question les notes de ce panéliste. Cependant, vous devez vous assurer de souligner aux panélistes que les commentaires de l'évaluation ne seront pas largement partagés ou ne refléteront pas leur participation à l'atelier. Ils sont donc fortement encouragés à partager leurs commentaires de façon honnête. Ces informations permettront d'améliorer les ateliers à venir.

Figure 35 : Conseils aux animateurs pour la présentation du formulaire d'évaluation

Les animateurs principaux et l'analyste de données compileront les notes d'évaluation après l'atelier. Les notations se présentent principalement sous la forme d'échelles de Likert, avec certaines zones de réponses ouvertes. Vous fournirez les résultats dans la documentation technique après l'atelier.

20. Clôture et logistique de l'atelier

MATÉRIEL : DIAPOSITIVES DE FACILITATION (PRÉSENTATION 20), DOSSIERS D'ATELIER À L'INTENTION DES PANÉLISTES

Lors de cette dernière séance d'atelier, encouragez les responsables gouvernementaux, les responsables de l'éducation des donateurs (le cas échéant) et les représentants des partenaires de mise en œuvre (le cas échéant) à formuler leurs remarques finales. Distribuez des certificats aux panélistes et remerciez-les pour leur participation (voir **l'annexe U** pour un modèle de certificat). Effectuez toute la logistique finale et prenez une photo de groupe, le cas échéant.

Figure 36 : Conseils aux animateurs pour la clôture de l'atelier

Les responsables doivent être encouragés à parler des prochaines étapes avec les critères de référence, c'est-à-dire l'utilisation de pourcentages par catégorie pour les rapports mondiaux. Si l'évaluation s'avérait être une évaluation basée sur un échantillon plutôt qu'un recensement, des travaux supplémentaires sur l'utilisation des poids d'échantillonnage peuvent être nécessaires en vue de la généraliser à la population.

G. CONSEILS SUPPLÉMENTAIRES POUR ORGANISER DES ATELIERS À DISTANCE

Voici des conseils pour organiser des ateliers à distance.

Logistique

- Assurez-vous que les panélistes disposent des documents imprimés dont ils auront besoin pour réaliser l'atelier.
- Assurez-vous que les panélistes peuvent y assister par le biais d'un ordinateur portable (fortement préféré) ou un smartphone afin qu'ils puissent voir les diapositives et soumettre des tâches. Autorisez les panélistes à soumettre des tâches soit sous forme de copies électroniques, de photos ou de numérisations de formulaires, ou (selon la tâche) dans le corps du texte par e-mail ou WhatsApp pour garantir que les panélistes sont en mesure d'effectuer des tâches avec des défis informatiques limités.
- Fournissez des cartes de données aux panélistes pour vous assurer qu'ils disposent de suffisamment de données pour se connecter aux sessions, et encouragez les panélistes à évaluer leur service bien avant l'atelier au cas où ils auraient besoin d'envisager de changer de fournisseur (si possible).
- Créez un groupe WhatsApp avant l'atelier pour faciliter les annonces, rappeler les sessions aux panélistes et assurer la facilité de communication entre les sessions de l'atelier lorsque de nombreux panélistes n'ont pas un accès régulier aux communications par courrier électronique.
- Envoyez des invitations de calendrier à tous les panélistes pour les sessions.
- Utilisez une plateforme de téléconférence qui permet : 1) de présenter des diapositives et de partager son écran, 2) d'affecter les panélistes à des groupes de discussion, 3) d'enregistrer les sessions (pour les panélistes qui manquent des parties de l'atelier en raison de problèmes technologiques pour les écouter après les sessions ; si possible, trouvez une plate-forme qui ne met pas de temps à traiter l'enregistrement afin qu'il puisse être diffusé rapidement aux panélistes), 4) couper le son de tout le monde à l'entrée dans la réunion, 5) produire des discussions dactylographiées, 6) lever la main pour indiquer une question ou commentaire et enregistrement des participants pour faciliter le suivi de la participation (si cela n'est pas possible, le personnel administratif doit être disponible pour suivre l'évolution de la participation tout au long de chaque session – en notant éventuellement qui est présent au début, au milieu et à la fin ; cela permet aux animateurs d'assurer un suivi auprès des panélistes qui ont manqué une partie importante de l'atelier en raison de problèmes technologiques).
- Organisez une série de courts appels préalables à l'atelier pour vérifier la capacité de petits groupes de panélistes à se connecter et à résoudre tout problème technologique.
- Demandez à un assistant administratif (NON PAS à un animateur) de gérer la plateforme de téléconférence, en laissant entrer les participants, en répartissant les panélistes en petits groupes, etc., car cette tâche peut être assez difficile à gérer lors de la conduite de sessions.
- Si vous utilisez des salles en petits groupes, fournissez à chaque panéliste des arrière-plans qui correspondent à son groupe afin d'identifier facilement s'il se trouve dans la mauvaise salle

Animateurs principaux

- Engagez deux (ou au moins un par niveau/matière/langue d'évaluation) animateurs principaux pour aider à animer les séances en petits groupes, pour permettre aux panélistes d'entendre plus d'une personne et pour permettre à une personne de suivre les questions qui surviennent dans le chat pendant que l'autre animateur assure la présentation.

Formation et interaction des animateurs de contenu

- Prévoyez au minimum une formation d'animateur de contenu à distance d'une durée de 8 heures, divisée en deux sessions. Toutefois, nous vous recommandons, si possible, d'augmenter la durée de cette formation pour garantir que les animateurs de contenu aient le temps de réaliser eux-mêmes chacune des activités.
- Demandez aux animateurs principaux de diriger toutes les séances plénières, à moins que les animateurs de contenu n'aient une expérience préalable en matière d'établissement de normes.
- En plus de la formation générale des animateurs de contenu, il est fortement recommandé de planifier de courtes séances de préparation avec les animateurs de contenu pour leur rappeler les questions clés juste avant les sessions au cours desquelles ils dirigent des groupes de travail.

Pré-sessions

Les ateliers à distance présentent l'avantage de pouvoir être prolongés sur une période un peu plus longue puisque les équipes de projet n'ont pas besoin de se préoccuper des modalités en matière d'hôtel et d'indemnités journalières (à moins que les panélistes ne se réunissent en présentiel avec uniquement les animateurs principaux présents à distance).

- Planifiez des pré-sessions pour permettre aux panélistes de se familiariser davantage avec le CM et l'évaluation avant d'entreprendre la tâche d'évaluation de l'apprenant avec trois apprenants qui répondent aux exigences de chaque GPL.
- Notez que dans certains cas, il peut être impossible pour les panélistes de terminer la tâche d'évaluation de l'apprenant (p. ex, en raison de problèmes de sécurité liés à la COVID-19). Dans ces cas-là, assurez-vous que les panélistes ont la possibilité de passer eux-mêmes l'évaluation au cours de l'une des pré-sessions ou d'administrer l'évaluation aux enfants dans leur foyer ou leur communauté (p. ex., à l'extérieur en utilisant des masques) entre les pré-sessions et la session ordinaire.
- Pour faciliter les tâches ultérieures, demandez aux panélistes d'écrire les noms des apprenants de leur classe qui sont décrits par les GPD « atteint » dans le cadre de leur activité entre les sessions.

Discussions

Un inconvénient majeur des ateliers à distance est que les panélistes n'ont pas la possibilité de s'engager dans des discussions informelles avec leurs voisins, ce qui met souvent en évidence des malentendus ou des questions, et les animateurs n'ont pas non plus la possibilité de se promener pendant que les panélistes accomplissent leurs tâches et regardent par-dessus leurs épaules pour identifier les malentendus potentiels. Les conseils ci-dessous visent à tenter de remédier à ces lacunes.

- Enregistrez chaque session et mettez les enregistrements à la disposition des panélistes pour qu'ils puissent les examiner en cas de problèmes de connexion
- Si possible, il serait intéressant d'identifier un moyen de permettre aux panélistes d'échanger verbalement entre eux, puis de revenir ensemble pour poser des questions aux animateurs. Cela peut être fait en se réunissant en petits groupes pendant 10 minutes après chaque série de diapositives afin de discuter et d'identifier les questions ou les problèmes. Il peut être nécessaire de prolonger les sessions pour tenir compte de cette possibilité.
- Si possible, il serait également intéressant d'identifier un moyen de regarder par-dessus les épaules des panélistes. Cela pourrait être fait en programmant des sessions individuelles de 15 à 30 minutes entre un animateur principal et chaque panéliste après la fin des séances plénières. Au

cours de ces appels, les animateurs peuvent demander aux panélistes d'expliquer la tâche et de décrire comment ils alignent/mettent en correspondance/notent chaque item. Cela devrait aider à identifier et à corriger les malentendus. Cela devrait également garantir que les panélistes qui ont manqué une partie de l'atelier en raison de problèmes technologiques aient le temps de poser des questions et de clarifier leur tâche.

- Enfin, les animateurs principaux peuvent rester en ligne pour chaque session d'atelier qui comprend une tâche assignée (tâche 1 et 3, pour les deux séries) pendant environ une heure après la session pour permettre aux personnes d'effectuer la tâche par elles-mêmes mais de rejoindre l'appel s'ils ont des questions.

CHAPITRE VI

CHAPITRE VI. AUTO-ÉVALUATION DES RÉSULTATS DE L'ATELIER

A. PRODUCTION DE LA DOCUMENTATION TECHNIQUE (UNE FOIS L'ATELIER TERMINÉ)

À la fin de l'atelier, les informations suivantes doivent être documentées à l'aide du modèle figurant à l'annexe V :

- Les données démographiques des panélistes
- Le(s) critère(s) de référence convenu(s) par l'atelier suite à la suppression des juges atypiques
- L'identification de tous les GPL qui semblent être affectés par des effets de plafond ou de plancher
- La proportion d'apprenants atteignant chacune des GPL pour lesquelles des critères de référence ont été fixés
- La précision, l'exactitude et la cohérence des jugements
- Les conclusions de l'évaluation des panélistes
- L'auto-évaluation des résultats au regard des critères #

Données démographiques des panélistes

Les données démographiques des panélistes auront été saisies à l'aide du formulaire de l'annexe M. L'équipe du projet devra confirmer que tous les panélistes satisfaisaient aux exigences de participation et, en tant que groupe, étaient suffisamment représentatifs. Les exigences relatives aux panélistes et les exigences de représentation sont fournies au chapitre IV.

Critères de référence

Au cours de l'atelier, des critères de référence ont été créés à partir des notes de tous les panélistes. Il est toutefois possible qu'un (ou plusieurs) panélistes soit considéré comme une valeur aberrante qui affecte de manière disproportionnée et inappropriée le critère de référence final. Les valeurs aberrantes sont déterminées à l'aide du modèle des clôtures de Tukey (Tukey, 1977) décrit à l'annexe J. Une fois les valeurs aberrantes supprimées, les références finales sont calculées comme la moyenne des panélistes restants.

Effets de plafond et de plancher

Un effet plafond se produit lorsqu'il n'y a pas suffisamment d'items difficiles dans une évaluation pour établir le critère de référence le plus élevé et se caractérise par un critère de référence proche de la note maximale. Un effet plancher se produit lorsqu'il n'y a pas suffisamment d'items faciles dans une évaluation pour définir le critère de référence le plus bas et se caractérise par un critère de référence proche de 0. Dans les deux cas, la définition du critère de référence est inappropriée car les informations issues de l'évaluation sont insuffisantes sur les performances probables des apprenants à ce niveau. Les critères d'alignement de la première auto-évaluation, garantissant qu'il y a suffisamment d'items alignés sur chacun des GPL, devraient éviter ce problème.

Shulruf (2016) a utilisé une étude de simulation pour examiner la fiabilité de la méthode Angoff. En interprétant cette étude, si le critère de référence est fixé à moins de 5 points du score minimum ou maximum, il est probable qu'il comporte trop d'erreurs pour être considéré comme fiable. En conséquence, les critères qui se situent à moins de 5 points de la note minimale ou maximale d'une évaluation ne seront pas acceptés pour établir des rapports au regard de l'ODD 4.1.1.

Données sur les résultats

Les critères finaux doivent être utilisés, ainsi que les distributions de données (calculées pour fournir des données d'impact au cours de l'atelier en utilisant l'**annexe N**) pour déterminer la proportion d'apprenants atteignant chacun des GPL.

Précision, exactitude et cohérence

Les formules de l'**annexe J** doivent être utilisées pour calculer les statistiques suivantes :

- Cohérence entre évaluateurs, avec suppression des panélistes marginaux
- L'erreur type pour chaque critère de référence
- Les intervalles de confiance pour chaque critère de référence

Bien qu'une valeur minimale soit fixée pour la cohérence entre évaluateurs (les valeurs égales ou supérieures à 0,7 sont considérées comme acceptables), il n'est pas possible de définir des valeurs similaires pour l'erreur type ou les intervalles de confiance car elles dépendent du nombre d'items de l'évaluation. Ceux-ci devraient être calculés et examinés par les pays eux-mêmes afin de déterminer s'ils sont appropriés pour l'évaluation. Les résultats des projets pilotes précédents montrent que pour une évaluation comportant environ 45 points de score, il est possible d'obtenir une erreur type inférieure à 1.

Évaluation

L'évaluation (**Annexe R**) contient une série d'énoncés sur lesquels les panélistes notent leur niveau d'accord sur une échelle de Likert à 5 points (1 – fortement en désaccord à 5 – fortement d'accord) regroupés en sections :

- Formation sur le GPF
- Formation à l'évaluation
- Tâche d'alignement
- Tâche de mise en correspondance
- Formation sur l'établissement de liens entre les politiques
- Résultats de la série 2

Il existe également une évaluation globale décrite sur une échelle de 4 points (1 – très mal à l'aise à 4 – très à l'aise).

Pour chaque affirmation, la note moyenne de tous les panélistes (à l'exclusion des valeurs aberrantes) doit être calculée. Pour chaque section, les notes minimales et maximales des énoncés doivent être notées. Pour l'évaluation globale, la note moyenne de tous les panélistes doit être calculée. Lorsque plusieurs ateliers pour différents niveaux/matières sont organisés simultanément, les résultats de l'évaluation doivent être calculés séparément pour chaque atelier.

Critères d'auto-évaluation

Les informations documentées ci-dessus doivent être utilisées pour confirmer que les résultats de l'atelier répondent aux exigences d'établissement de rapports au regard de l'ODD 4.1.1.

- Critère 1 : tous les panélistes satisfaisaient-ils aux conditions de participation ? (OUI / NON)
- Critère 2 : le groupe de panélistes était-il suffisamment représentatif au regard des caractéristiques convenues par le pays ? (OUI / NON)
- Critère 3 : toutes les valeurs aberrantes ont-elles été supprimées avant de calculer les critères finaux ? (OUI / NON)
- Critère 4 : des critères de référence ont-ils été définis que pour des GPL qui ne présentent pas d'effets de plancher ou de plafond ? (OUI / NON)
- Critère 5 : la statistique de cohérence inter-évaluateurs est-elle supérieure ou égale à 0,7 ? (OUI / NON)
- Critère 6 : l'erreur type pour chaque critère de référence a-t-elle été calculée et examinée pour être déterminée de manière appropriée ? (OUI / NON)

- Critère 7 : l'intervalle de confiance pour chaque critère de référence a-t-il été calculé et examiné pour être déterminé de manière appropriée ? (OUI / NON)
- Critère 8 : la note moyenne pour chaque section de l'évaluation était-elle supérieure ou égale à 4 ? (OUI/ NON)
- Critère 9 : la moyenne des notes de l'évaluation globale était-elle supérieure ou égale à 3 ? (OUI / NON)

Afin d'être considérés comme éligibles pour le rapport au regard de l'ODD 4.1.1, les pays devront être en mesure de répondre OUI à chacune des questions critères. Si la réponse à une ou plusieurs questions est NON, les pays devront alors envisager d'organiser à nouveau l'atelier ou d'envisager d'autres alternatives en matière de reporting.

B. SOUMETTRE DES PREUVES À L'ISU

Le tableau 12 indique les informations que les pays devront fournir à l'ISU pour les rapports sur l'ODD 4.1.1.

Tableau 12 : informations requises pour rendre compte au regard de l'ODD 4.1.1

Niveau	Identifiant des indicateurs	Description des indicateurs
ODD 4.1.1a	MATH.G2T3	Proportion d'élèves (filles et garçons) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en mathématiques (%)
	MATH.G2T3.F	Proportion d'élèves (filles) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en mathématiques (%)
	MATH.G2T3.G	Proportion d'élèves (garçons) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en mathématiques (%)
	LECTURE.G2T3	Proportion d'élèves (filles et garçons) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en lecture (%)
	LECTURE.G2T3.F	Proportion d'élèves (filles) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en lecture (%)
	LECTURE.G2T3.G	Proportion d'élèves (garçons) de 2 ^{ème} ou 3 ^{ème} année atteignant au moins un niveau minimal de compétence en lecture (%)
ODD 4.1.1b	MATH.PRIMAIRE	Proportion d'élèves (filles et garçons) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
	MATH.PRIMAIRE.F	Proportion d'élèves (filles) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
	MATH.PRIMAIRE.G	Proportion d'élèves (garçons) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
	LECTURE.PRIMAIRE	Proportion d'élèves (filles et garçons) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en lecture (%)
	LECTURE.PRIMAIRE.F	Proportion d'élèves (filles) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en lecture (%)
	LECTURE.PRIMAIRE.G	Proportion d'élèves (garçons) à la fin de l'enseignement primaire atteignant au moins un niveau minimal de compétences en lecture (%)
ODD 4.1.1c	MATH. FIN DE	Proportion d'élèves (filles et garçons) à la fin de l'enseignement

L'ENSEIGNEMENT SECONDAIRE	secondaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
MATH. FIN DE L'ENSEIGNEMENT SECONDAIRE. F	Proportion d'élèves (filles) à la fin de l'enseignement secondaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
MATH. FIN DE L'ENSEIGNEMENT SECONDAIRE. G	Proportion d'élèves (garçons) à la fin de l'enseignement secondaire atteignant au moins un niveau minimal de compétences en mathématiques (%)
LECTURE. FIN DE L'ENSEIGNEMENT SECONDAIRE	Proportion d'élèves (filles et garçons) à la fin de l'enseignement secondaire atteignant au moins un niveau minimal de compétences en lecture (%)
LECTURE. FIN DE L'ENSEIGNEMENT SECONDAIRE. F	Proportion d'élèves (filles) à la fin de l'enseignement secondaire atteignant au moins un niveau minimal de compétences en lecture (%)
LECTURE. FIN DE L'ENSEIGNEMENT SECONDAIRE. G	Proportion d'élèves (garçons) à la fin de l'enseignement secondaire atteignant au moins un niveau minimal de compétences en lecture (%)

Les pays devront soumettre à l'ISU les preuves suivantes pour démontrer que les informations du **tableau 12** ont été générées conformément aux exigences de l'établissement de liens entre les politiques :

- Auto-évaluation – pertinence de l'évaluation (consulter l'**annexe C**)
- Auto-évaluation – résultats de l'atelier (consulter l'**annexe V**)
- Rapport de l'atelier consacré à l'établissement de liens entre les politiques (voir l'**annexe V**)

BIBLIOGRAPHIE

BIBLIOGRAPHIE

- Adams, R., Jackson, J., & Turner, R. (2018). *Learning progressions as an inclusive solution to global education monitoring*. Melbourne, Australia: Australian Council for Educational Research.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thondike (Ed.) *Educational Measurement* (2nd ed.). Washington, DC.: American Council on Education.
- Bejar, I. (1983). Subject matter experts' assessment of item statistics. *Applied Psychological Measurement*, 7(3), 303-310.
- Berk, R. (1996). Standard setting: The next generation (where few psychometricians have gone before!). *Applied Measurement in Education*, 9(3), 215-225.
- Brennan, R. L. (2004). *BB-CLASS v.1.1 [Computer program]*. Iowa City: Center for Advanced Studies in Measurement and Assessment, University of Iowa.
- Brennan, R. L., & Wan, L. (2004). Bootstrap procedures for estimating decision consistency for single administration complex assessments. *CASMA Research Report No. 7*. Iowa City: University of Iowa.
- Brown, J.D. (1989). Criterion-referenced test reliability. *University of Hawai'i Working Papers in ESL*, 8(1), 79-113.
- Burry-Stock, J.A., Shaw, D.G., Laurie, C., & Chissom, B.S. (1996). Reader agreement indexes for performance assessments. *Educational and Psychological Measurement*, 56, 251-262.
- Chang, L. (1999). Judgmental item analysis of the Nedelsky and Angoff standard-setting methods. *Applied Measurement in Education*, 12(2), 151-165.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37- 46.
- Engelhard, G. & Stone, G. (1998). Evaluating the quality of ratings obtained from standard-setting judges. *Educational and Psychological Measurement*, 58(2), 179-196.
- Ferdous, A. (2019). *Setting performance standards for reading fluency in Lebanon*. Paper presented for the annual meeting of the Comparative and International Education Society. San Francisco, CA.
- Ferdous, A. & Buckendahl, C. (2013). Evaluating panelists' standard setting perceptions in a developing nation. *International Journal of Testing*, 13(1), 4-18.
- Ferdous, A. & Plake, B. (2005). Understanding the factors that influence decisions of panelists in a standard setting study. *Applied Measurement in Education*, 18(3), 257-267.
- Ferdous, A. & Plake, B. (2007). A mathematical formulation for computing inter-panelist inconsistency for Body of Work, Bookmark, and Yes/No Variation of Angoff methods. Paper presented at the annual meeting of the National Council on Measurement in Education (NCME), Chicago, IL.

- Fleiss, J. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382.
- Frisbie, D.A. (2003). Checking the alignment of an assessment tool and a set of content standards. Iowa City, IA: University of Iowa.
- Giraud, G., Impara, C., & Plake, B. (2005). Teachers' conceptions of the target examinee in Angoff standard setting. *Applied Measurement in Education*, 18(3), 223-232.
- Halpin, G. & Halpin, G. (1983). *Reliability and validity of ten different standard setting procedures*. Paper presented at the American Psychological Association, Anaheim, CA.
- Hambleton, R. (2001). The next generation of the ITC Test Translation and Adaptation Guidelines. *European Journal of Psychological Assessment*, 17(3), 164-172.
- Hambleton, R. (2008). Psychometric models, test designs and item types for the next generation of educational and psychological tests. In D. Bartram and R. Hambleton (Eds.) *Computer-Based Testing and the Internet: Issues and Advances* (pp. 77-89). New York, NY: John Wiley & Sons Ltd.
- Hambleton, R. & Bourque, M. (1991). The LEVELS of mathematics achievement: Initial performance standards for the 1990 NAEP mathematics assessment: Vol. III. Technical report. Washington, D.C.: National Assessment Governing Board.
- Hambleton, R. & Pitoniak, M. (2006). Setting performance standards. In R. L. Brennan (Ed.) *Educational Measurement* (4th ed.). Westport: American Council on Education & Praeger Publishers.
- Hambleton, R. & Plake, B. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education*, 8(1), 41-55.
- Hanson, B. A., & Brennan, R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. *Journal of Educational Measurement*, 27, 345-359.
- Hurtz, G. & Hertz, N. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability theory study. *Educational and Psychological Measurement*, 59(6), 885-897.
- Jaeger, R. (1989). Certification of learner competence. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Jaeger, R. (1995). Setting performance standard through two-stage judgmental policy capturing. *Applied Measurement in Education*, 8(1), 15-40.
- Kahl, S., Crockett, T., DePascale, C., & Rindfleisch, S. (1995). *Setting standards for performance levels using learner-based constructed-response method*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Kane, M. (1998). Choosing between examinee-centered and test-centered standard-setting methods. *Educational Assessment*, 5(3), 129-145.

- Lewis, D., Mitzel, H., Green, D., & Patz, R. (1999). *The bookmark standard setting procedure*. Monterey, CA: McGraw- Hill.
- Livingston, S. & Zieky, M. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Princeton, NJ: Educational Testing Service.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179-197.
- Lorge, I. & Kruglov, L. (1953). The improvement of estimates of test difficulty. *Educational and Psychological Measurement*, 13(1), 34-46.
- Management Systems International (2019). *Policy linking method: Linking assessments to a global standard*. U.S. Agency for International Development (USAID), Washington, D.C.
- Mehrens, W. & Popham, W. (1992). How to evaluate the legal defensibility of high-stakes tests. *Applied Measurement in Education*, 5(3), 265-283.
- Norcini, J., Shea, J., & Grasso, L. (1991). The effect of numbers of experts and common items on cutting score equivalents based on expert judgement. *Applied Psychological Measurement*, 15(3), 241-246.
- Pitoniak, M. (2003). Standard setting methods for complex licensure examinations. *Doctoral Dissertations 1896 – February 2014*. University of Massachusetts, Amherst.
- Plake, B., Ferdous, A., & Buckendahl, C. (2005). *Setting multiple performance standards using the yes/no method: An alternative item mapping method*. Paper presented to the meeting of the National Council on Measurement in Education (NCME), Montreal, Canada.
- Plake, B. & Hambleton, R. (2000). A standard setting method designed for complex performance assessments: Categorical assignments of learner work. *Educational Assessment*, 6(3), 197-215.
- Plake, B., Hambleton, R., & Jaeger, R. (1997). A new standard-setting method for performance assessments: The dominant profile judgment method and some field-test results. *Educational and Psychological Measurement*, 57(3), 400- 411.
- Rasch, G. (1960) Probabilistic models for some intelligence and attainment tests. Copenhagen: Danmarks Paedagogiske Institut.
- Schaeffer, G. & Collins, J. (1984). *Setting performance standards for high-stakes tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Shulruf, B., Wilkinson, T., Weller, J. et al. (2016) Insights into the Angoff method: results from a simulation study. *BMC Med Educ* 16, 134.
- Subkoviak, M. J. (1976). Estimating reliability from a single administration of a criterion-referenced test. *Journal of Educational Measurement*, 13, 265-276.
- Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability for mastery tests. *Journal of Educational Measurement*, 25, 47-55.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Massachusetts.

UNESCO. (2018a). *Global content framework of reference for reading: Global consultation*. Paper presented at the fifth meeting of the Global Alliance to Monitor Learning (GAML), Hamburg, Germany.

UNESCO. (2018b). *Global content framework of reference for mathematics: Global consultation*. Paper presented at fifth meeting of the Global Alliance to Monitor Learning (GAML), Hamburg, Germany.

ANNEXES

ANNEXE A – RESSOURCES CONNEXES

- [Cadre mondial de compétences en mathématiques : 1^{ère} à la 9^{ème} année](#)²²
- [Cadre mondial de compétences en lecture : 1^{ère} à la 9^{ème} année](#)²³
- Diapositives de l’animation de l’atelier : établissement de liens entre les politiques pour mesurer les résultats d’apprentissage mondiaux et les évaluations chronométrées
- Diapositives de l’animation de l’atelier : liens établissement entre les politiques pour mesurer les résultats d’apprentissage mondiaux et les évaluations non chronométrées
- Diapositives de l’animateur de contenu
- Liste de contrôle de préparation à l’atelier
- Formulaire d’évaluation d’alignement pour la tâche I
- Formulaire d’évaluation des items
- Formulaire de renseignements démographiques sur les panélistes
- Formulaire récapitulatifs des modèles d’auto-évaluation
- Modèles
 - Lettre d’invitation aux observateurs
 - Lettre d’invitation pour les panélistes de l’atelier
 - Certificat d’appréciation

²² <https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2021/03/Global-Proficiency-Framework-Math.pdf>

²³ <https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2021/03/Global-Proficiency-Framework-Reading.pdf>

ANNEXE B – NIVEAU MINIMAL DE COMPÉTENCES MONDIAL

En deçà du niveau Partiellement atteint les compétences minimales : les apprenants ne disposent pas des connaissances et des compétences les plus élémentaires. En conséquence, ils ne peuvent généralement pas accomplir les tâches les plus élémentaires.

Partiellement atteint les compétences minimales : les apprenants ont des connaissances et des compétences partielles. En conséquence, ils peuvent accomplir partiellement les tâches de base.

Atteint les compétences minimales : les apprenants possèdent des connaissances et des compétences suffisantes. En conséquence, ils peuvent accomplir avec succès des tâches de base.

Dépasse les compétences minimales : les apprenants ont des connaissances et des compétences supérieures. En conséquence, ils peuvent accomplir avec succès des tâches complexes.

Une description plus complète des MPL se trouve [dans le](#)²⁴ document MPLs Unpacked.

²⁴ https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2022/11/WG_GAML_4_MPLs-Unpacked_ACER.pdf

ANNEXE C : RÉSUMÉ DU MODÈLE D'AUTO-ÉVALUATION (PERTINENCE DE L'ÉVALUATION)

Outil d'évaluation	[Insérer le nom de l'outil]
Juridiction	[Insérer la juridiction dans laquelle l'outil d'évaluation est administré]
Niveau	[Insérer la note évaluée par l'outil]
Niveau ODD 4.1.1	a / b / c [rayer la mention inutile]
Sujet	Mathématiques / Lecture [rayer la mention inutile]
GPL en cours de définition	Partiellement atteint / Atteint / Dépasse [rayer la mention inutile]
Date de l'auto-évaluation	[Insérer la date à laquelle l'auto-évaluation a été réalisée]

Critère 1 – Alignement

Évaluateurs	[Insérer les noms et les organisations de ceux qui ont entrepris l'exercice d'alignement]
Niveau d'alignement	Minime / Supplémentaire / Fort [rayer la mention inutile]
Nombre de points contenus dans l'outil d'évaluation	[Insérer le nombre de points]
Nombre de points par domaine concerné	[Insérer le nombre de points par domaine concerné pour le niveau d'alignement]
Nombre de sous-concepts dans les domaines concernés	[Insérer le nombre de sous-concepts dans les domaines concernés pour le niveau d'alignement]
Nombre de sous-concepts pertinents évalués	[Insérer le nombre de sous-concepts pertinents couverts par]
Pourcentage de sous-concepts pertinents évalués	[Insérer le pourcentage de sous-concepts pertinents évalués]
L'outil d'évaluation satisfait-il aux exigences quantitatives du cahier des charges ?	Oui / Non [rayer la mention inutile]
Les cadres d'évaluation et du programme d'études sont-ils alignés ?	Oui / Non [rayer la mention inutile]

Note du critère 1 : Insuffisant / Minimal / Bon / Excellent [rayer la mention inutile]

Critère 2 – Examen de l'item

Évaluateurs	[Insérer les noms et les organisations de ceux qui ont entrepris l'auto-évaluation]
Existe-t-il des preuves que les items de l'évaluation ont été examinés quantitativement ?	Oui / Non [rayer la mention inutile]
Existe-t-il des preuves que les items de l'évaluation ont été examinés qualitativement ?	Oui / Non [rayer la mention inutile]
Existe-t-il des preuves que les items ont été examinés pour garantir leur pertinence pour les sous-groupes de la population concernés ?	Oui / Non [rayer la mention inutile]

Note du critère 2 : Insuffisant / Minimal / Excellent [rayer la mention inutile]

Critère 3 – Échantillon

Évaluateurs	[Insérer les noms et les organisations de ceux qui ont entrepris l'auto-évaluation]
L'évaluation a-t-elle été administrée à l'ensemble de la cohorte ou à un échantillon ?	Cohorte entière/échantillon [rayer la mention inutile]
Certains sous-groupes de la population ont-ils été systématiquement exclus de l'administration ?	[Insérer les sous-groupes exclus de la population à des fins de déclaration]
Pour les évaluations basées sur un échantillon, la marge d'erreur est-elle de 5 % ou moins au niveau de	Oui / Non / Sans objet [rayer la mention inutile]
La taille minimale de l'effet détectable a-t-elle été calculée et réfléchie avant de finaliser les calculs de la taille de l'échantillon ?	Oui / Non / Sans objet [rayer la mention inutile]

Notation du critère 3 : Insuffisant / Minime / Excellent [rayer la mention inutile]

Critère 4 – Administration

Évaluateurs	[Insérer les noms et les organisations de ceux qui ont entrepris l'auto-évaluation]
L'outil d'évaluation a-t-il été administré de manière appropriée et normalisée ?	Oui / Non [rayer la mention inutile]
Les guides d'administration étaient-ils clairs sur le processus administratif ?	Oui / Non [rayer la mention inutile]
Les incidents significatifs d'administration inappropriée ont-ils été enregistrés et les résultats pertinents ont-ils été exclus des résultats ?	Oui / Non [rayer la mention inutile]
L'exclusion des résultats d'évaluations administrées de manière inappropriée a-t-elle affecté la représentativité de l'échantillon ?	Oui / Non / Sans objet [rayer la mention inutile]

Note du critère 4 : Insuffisant / Minime [rayer la mention inutile]

Critère 5 – Fiabilité

Évaluateurs	[Insérer les noms et les organisations de ceux qui ont entrepris l'auto-évaluation]
La valeur du coefficient alpha (ou statistique de fiabilité équivalente) pour l'évaluation est-elle supérieure ou égale à 0,7 ?	Oui / Non [rayer la mention inutile]
Existe-t-il des preuves de dispositions appropriées en matière d'assurance qualité pour tous les items notés par l'homme ?	Oui / Non / Sans objet [rayer la mention inutile]
Le niveau d'accord entre les évaluateurs humains et les scores préalablement convenus, ou ceux doublement notés, est-il supérieur à 80 % ?	Oui / Non / Sans objet [rayer la mention inutile]
Les autres mesures de fiabilité de l'évaluation, par exemple la constance de la classification, l'exactitude de la classification ou la fiabilité entre les évaluateurs, sont-elles à des niveaux conformes aux meilleures pratiques internationales ?	[Insérer les détails des autres mesures de fiabilité] Oui / Non / Sans objet [rayer la mention inutile]

Notation du critère 5 : Insuffisant / Minime / Bon / Excellent [rayer la mention inutile]

Note d'auto-évaluation globale

Critères	Insuffisant	Minime	Bon	Excellent
Critère 1 – Alignement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Critère 2 – Examen de l'item	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Critère 3 – Échantillon	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
Critère 4 – Administration	<input type="checkbox"/>	<input type="checkbox"/>		
Critère 5 – Fiabilité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Note d'auto-évaluation globale	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ANNEXE D – LISTE DE CONTRÔLE POUR LA PRÉPARATION DE L'ATELIER

Table 13 : liste de contrôle pour la préparation de l'atelier

Activité	Responsable	Échéance	✓	Commentaires
1. Sélectionner et recruter des animateurs et des analystes de données				
a. Identifier et recruter des animateurs principaux				
b. Identifier et recruter des animateurs de contenu				
c. Identifier et recruter un coordinateur, si nécessaire pour la préparation logistique				
2. Préparer la logistique de l'atelier				
a. Déterminer si l'atelier se déroulera en présentiel, mixte (panélistes et animateurs de contenu en personne et animateurs principaux à distance) ou à distance				
b. Identifier et sécuriser un espace physique ou un service de conférence à distance				
c. Déterminer quelle nourriture/quels rafraîchissements seront fournis aux participants et se les procurer				
d. Organiser ou se procurer du matériel, tel que des cahiers, des stylos, des tableaux à feuilles mobiles, des dossiers, des porte-noms/chapiteaux, des bannières				
e. Identifier les indemnités journalières, le budget de déplacement, les indemnités de carte téléphonique/de données (pour les ateliers à distance), les frais d'hôtel, etc., et convenir, avec le gouvernement/l'agence d'évaluation et les responsables des donateurs (le cas échéant), des montants pour les panélistes et les observateurs				
f. Prendre les dispositions nécessaires pour les chambres d'hôtel, si nécessaire				
g. Prendre les dispositions nécessaires pour les déplacements des animateurs, si nécessaire				
h. Prendre les dispositions nécessaires pour les déplacements des panélistes/observateurs, si nécessaire				
i. Inspecter le lieu pour planifier l'atelier et l'emplacement des salles en petits groupes				
j. Identifier la méthode de réception des fonds dans le pays (si nécessaire) ; cela peut impliquer un virement bancaire ou un paiement en espèces				
k. Effectuer un virement bancaire/un paiement en espèces, si nécessaire				
l. Transférer des fonds aux participants ; pour les ateliers en présentiel, cela se fait souvent pendant l'atelier				
3. Sélectionner et inviter des participants				
a. Finaliser la liste des enseignants panélistes				
b. Finaliser la liste des panélistes spécialistes du programme d'études				
c. Finaliser la liste des observateurs				
Rédiger des instructions pour l'activité d'évaluation préalable à l'atelier, si celui-ci se tient en présentiel				
e. Préparer une évaluation de la pratique si la sécurité de l'évaluation est une préoccupation (voir la figure 13 pour plus d'informations)				
f. Préparer et distribuer des invitations, avec des instructions d'évaluation préalable à l'atelier, aux enseignants panélistes				
g. Préparer et distribuer des invitations, avec des instructions d'évaluation préalable à l'atelier, aux panélistes spécialisés				

Activité	Responsable	Échéance	✓	Commentaires
h. Préparer et distribuer les invitations aux observateurs				
4. Préparer les documents				
a. Finaliser et imprimer le programme (et le distribuer si l'atelier se déroule à distance)				
b. Finaliser et imprimer la liste des acronymes (et le distribuer si l'atelier se déroule à distance)				
c. Finaliser et imprimer le glossaire (et la distribuer si l'atelier se déroule à distance)				
d. Attribuer des identifiants aux panélistes (et les distribuer si l'atelier se déroule à distance)				
e. Traduire le GPF de lecture dans la langue locale, si nécessaire				
f. Adapter le GPF aux niveaux/matières concernés, l'imprimer (et le distribuer si l'atelier se déroule à distance)				
g. Élaborer des passages/questions pratiques pour les diapositives				
h. Finaliser les formulaires d'évaluation (formulaires d'alignement et d'évaluation des items), les imprimer (et les distribuer si l'atelier se déroule à distance)				
i. Imprimer/distribuer les outils d'évaluation, conformément aux protocoles de sécurité				
j. Finaliser et imprimer les formulaires d'évaluation des ateliers				
k. Analyser les données pour produire des distributions de données, des données sur la difficulté des items, etc.				
l. Finaliser les diapositives de facilitation et les imprimer				
m. Finaliser les feuilles de présence quotidiennes et les imprimer				
5. Former les animateurs de contenu				
a. Finaliser les diapositives pour la formation des animateurs de contenu				
b. Prendre des dispositions logistiques pour la formation des animateurs de contenu				
c. Former les animateurs de contenu				

Coordinateur : _____

Animateur principal : _____

ANNEXE E – PLANIFICATEUR D'ACTIVITÉS D'ATELIER

Tableau 14 : Planificateur d'activités d'atelier

CALENDRIER SEMAINE PAR SEMAINE POUR LES ACTIVITÉS DE PRÉPARATION DES ATELIERS CONSACRÉS A L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES					
Nombre	Activité	Rôle/Responsabilité	Format de l'atelier pour lequel l'activité est appropriée	Échéance (nombre de semaines avant l'atelier)	
4 semaines avant l'atelier					
1	Prendre contact avec le pays	ISU/Organisation donatrice (DO)	Tous	Au moins 4 semaines	
2	Décider sur quelle évaluation, quel niveau scolaire et quelle langue vous concentrer	Pays bénéficiant du soutien de l'ISU/DO et du partenaire de mise en œuvre (DP)	Tous	4 semaines	
3	Décider du format que prendra l'atelier (en présentiel, en distanciel ou hybride avec les participants se réunissant dans un ou plusieurs lieux) et du calendrier de l'atelier	Pays bénéficiant du soutien de l'ISU/DO/DP	Tous	4 semaines	
4	Identifier les animateurs de contenu locaux	Pays	Tous	4 semaines	
5	Identifier les interprètes (le cas échéant)	Pays	Tous	4 semaines	
6	Identifier le logisticien (si nécessaire)	Pays	Tous	4 semaines	
7	Identifier les panélistes (enseignants et spécialistes du contenu), notamment en collectant leurs coordonnées : s'assurer que le panel est représentatif	Pays	Tous	4 semaines	
8	Adapter le GPF aux notes/matières concernées	DP	Tous	4 semaines	
3 à 4 semaines avant l'atelier (le plus tôt sera le mieux)					
9	Ordre du jour provisoire	DP	Tous	3 à 4 semaines	
10	Soumettre vos commentaires sur l'ordre du jour provisoire	Pays	Tous	3 à 4 semaines	
11	Finaliser l'ordre du jour	DP	Tous	3 à 4 semaines	
12	Inviter des panélistes	Pays, ISU/DO ou DP - selon les préférences du pays	Tous	3 à 4 semaines (en fonction des normes du pays)	
3 semaines avant l'atelier					
13	Identifier et inviter tous les observateurs de l'atelier - provenant d'autres donateurs, ministères, etc.	Pays bénéficiant du soutien de l'ISU/DO/DP	Tous	3 semaines	
14	Identifier les autres coûts potentiels pour l'atelier, y compris les cartes de téléphone/ Internet, le transport, l'hébergement, les indemnités journalières, les repas, l'eau et le matériel utilisé pendant l'atelier	Pays	Tous	3 semaines	
15	Soumettre le budget à l'ISU/DO	Pays	Tous	3 semaines	
16	L'ISU/DO et le DP finalisent les accords de non-divulagation	ISU/DO et DP	Tous	3 semaines	
Nombre	Activité	Rôle/Responsabilité	Format de l'atelier pour lequel l'activité	Échéance (nombre de semaines avant l'atelier)	

				est appropriée	
17	Envoyer les outils d'évaluation à l'ISU/DO et au DP	Pays	Tous	3 semaines	
18	Fournir le logo du ministère pour les certificats et la bannière (cette dernière uniquement pour les ateliers en présentiel et hybrides) et déterminer qui, au ministère, signera	Pays	Tous	3 semaines	
19	Traduire le GPF dans la langue locale, si nécessaire, et envoyer la traduction pour relecture pour en vérifier la qualité	Pays	Tous	3 semaines	
20	Rédiger des diapositives d'atelier, y compris des exemples d'items, et des formulaires d'évaluation à envoyer à l'ISU/DO pour examen	DP	Tous	3 semaines	
21	Identifier et sécuriser l'espace physique pour l'atelier	Pays	Hybride	3 semaines	
2 semaines avant l'atelier					
22	Finaliser les contrats avec les animateurs de contenu, les interprètes et le logisticien locaux (les deux derniers, le cas échéant)	ISU/DO	Tous	2 semaines	
23	Examiner les diapositives de l'atelier, notamment les exemples d'items et les formulaires d'évaluation, et envoyer vos commentaires au DP	ISU/DO	Tous	2 semaines	
24	Finaliser le protocole d'accord avec le pays en fonction du budget approuvé	ISU/DO	Tous	2 semaines	
25	Projets de certificats et bannière	DP	Tous	2 semaines	
26	Finaliser les formulaires et les diapositives d'évaluation des items en fonction des retours de l'ISU/DO	DP	Tous	2 semaines	
27	Envoyer les données à l'ISU/DO et au DO (si possible)	Pays	Tous	2 semaines	
28	Confirmer la participation des panélistes	Pays	Tous	2 semaines	
29	Réserver des chambres d'hôtel pour les panélistes, si nécessaire	Pays	En présentiel et hybride	2 semaines	
30	Choisir un service de conférence à distance pour l'atelier	Pays	À distance	2 semaines	
31	Transférer des fonds et/ou des cartes téléphoniques/internet aux participants	Pays	À distance	2 semaines	
32	Finaliser les diapositives pour la formation des animateurs de contenu	DP	Tous	2 semaines	
33	Prendre des dispositions logistiques pour la formation des animateurs de contenu	DP	Tous	2 semaines	
1 semaine avant l'atelier					
34	Préparer les fonds à distribuer aux participants pour les indemnités journalières, les déplacements, etc.	Pays	En présentiel et hybride	1 semaine	
35	Déterminer quelle nourriture/quels rafraîchissements seront fournis aux participants et se les procurer	Pays	En présentiel et hybride	1 semaine	
36	Organiser ou se procurer du matériel, tel que des cahiers, des stylos, des tableaux à feuilles mobiles, des dossiers, des porte-noms/chapiteaux, des bannières	Pays	En présentiel et hybride	1 semaine	
37	Inspecter le lieu pour planifier l'atelier, l'emplacement des salles de petits groupes et tester l'accès à distance (le cas échéant, p. ex. s'il ne s'agit pas d'un établissement gouvernemental)	Pays	En présentiel et hybride	1 semaine	
38	Finaliser le programme (avec les éventuels changements de dernière minute)	DP	Tous	1 semaine	
39	Finaliser le programme, la liste des acronymes, le glossaire, l'évaluation, le GPF, les formulaires de notation, les formulaires d'évaluation, les diapositives avec les champs de notes, les certificats, les bannières et tout autre document	DP	Tous	1 semaine	

Nombre	Activité	Rôle/Responsabilité	Format de l'atelier pour lequel l'activité est appropriée	Échéance (nombre de semaines avant l'atelier)
40	Imprimer le programme, la liste des acronymes, le glossaire, l'évaluation, le GPF, les formulaires de notation, les formulaires d'évaluation, les diapositives avec les champs de notes, les certificats, les bannières et tout autre document	Pays	Tous	1 semaine
41	Distribuer le programme, la liste des acronymes, le glossaire, l'évaluation, le GPF, les formulaires de notation, les formulaires d'évaluation, les diapositives avec les champs de notes, les certificats, les bannières et tout autre document	Pays	À distance	1 semaine
42	Attribuer des identifiants aux panélistes	DP	Tous	1 semaine
43	Distribuer des identifiants aux panélistes	Pays	À distance	1 semaine
44	Analyser les données pour produire des distributions de données, des données sur la difficulté des items, etc.	DP	Tous	1 semaine
45	Finaliser les diapositives de facilitation et les imprimer	DP	En présentiel et	1 semaine
46	Finaliser les feuilles de présence quotidiennes	DP	Tous	1 semaine
47	Imprimer les feuilles de présence quotidiennes	Pays	En présentiel et	1 semaine
48	Former les animateurs de contenu	DP	Tous	1 semaine
Quelques jours avant l'atelier				
49	Test de la plateforme à distance avec des panélistes ou un lieu pour s'assurer que les participants peuvent accéder à la plateforme et n'ont pas besoin d'assistance technique	Tous	Tous	Quelques jours
Début de l'atelier				

ANNEXE F – MODÈLE D'ESTIMATION DU BUDGET

MODÈLE D'ESTIMATION DES COÛTS POUR L'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES					
Date/heure proposée :					
Sl. N°	Activités	Nb de personnes	Nb de jours	Coût unitaire	Budget estimé
Coûts du temps du personnel et des consultants					
1	Animateurs de contenu				
2	Logisticien, si cela s'avère nécessaire				
3	Interprètes, si cela s'avère nécessaire				
Coûts de participation du personnel et des consultants					
4	Indemnité journalière de subsistance (IJS) pour les participants/animateurs/interprètes et				
5	Participants rémunérés				
6	animateurs rémunérés				
7	Transport des participants/animateurs/interprètes				
8	Chambres d'hôtel (si l'atelier se déroule en				
9	Kit d'atelier (comprend dossier, bloc-notes, stylo, crayon, gomme, taille-crayon, etc.)				
10	Déjeuner si besoin				
11	Thé/café, collations et eau				
Coûts logistiques préalables à l'atelier					
12	Impression et livraison (envoi) des invitations/nominations, si nécessaire				
13	Traduction du GPF dans la langue locale				
14	Traduction de tout autre document dans la langue				
Lieu de l'atelier et coûts des matériels					
15	Location de salle, si nécessaire				
16	Impression des diapositives, de l'ordre du jour, du GPF, des évaluations, des formulaires de notation, de la liste des acronymes, du glossaires, des feuilles de présence et des attestations				
17	Impression de bannière				
18	Location de matériel de visioconférence				
19	Communications et Internet (forfait)				
20	Soutien administratif (3 %)				
Budget total					-

ANNEXE G – DIAPOSITIVES DE FACILITATION DE L’ATELIER

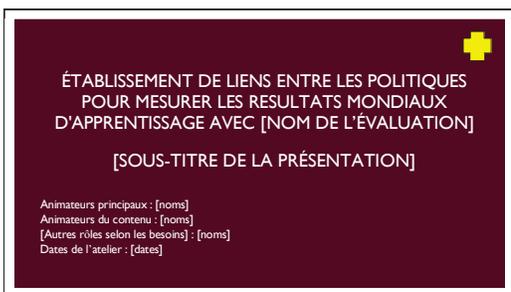
Les animateurs devront mettre à jour/adapter toutes les diapositives marquées d’un signe « plus » jaune  pour les utiliser dans leur contexte spécifique. Les instructions sur la manière de le faire sont incluses en **GRAS** dans la section des notes de chaque diapositive. Veuillez retirer le signe « plus » jaune avant d’utiliser la présentation.

Des notes pour les animateurs sont également incluses dans la section des notes.

Des crochets, comme ceux-ci [], ont été utilisés pour désigner les zones nécessitant une mise à jour/adaptation sur les diapositives réelles.

PRÉSENTATION 1 – ACCUEIL ET PRÉSENTATIONS

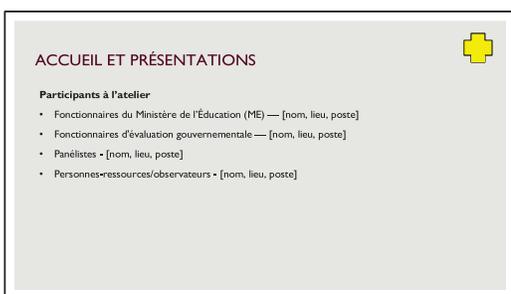
Slide 1



Slide 2



Slide 3



Slide 4

ACCUEIL ET PRÉSENTATIONS 

Équipe du projet

- [Donateur, le cas échéant] fonctionnaires de l'éducation — [nom, poste]
- Représentants du partenaire de mise en œuvre (PMO, le cas échéant) — [nom, poste]
- Coordinateur(s) de l'atelier [nom, poste]
- Animateur(s) principal(aux) [nom, poste]
- Animateurs de contenu (groupe) [nom, poste]
- Personnel administratif [nom, poste]

Slide 5

[AJOUTER LES DIAPOSITIVES NECESSAIRES POUR TOUTES LES PERSONNALITES FAISANT DES REMARQUES D'OUVERTURE] 

Slide 6

APERÇU DE L'ATELIER 

- [Insérer un aperçu de la logistique de l'atelier, par exemple, le nombre de jours/le nombre de séances, les heures de début et de fin, etc.]
- L'atelier comprendra **des présentations par les animateurs** ainsi que **des activités** que les panélistes devront réaliser en groupes.
- **Nous aborderons trois tâches principales** au cours de l'atelier.

Slide 7

OBJECTIFS DE L'ATELIER 

À la fin de cet atelier, notre objectif est de :

- comprendre dans quelle mesure le [nom de l'évaluation] est en accord avec les attentes mondiales en [matières] pour [niveaux scolaires]
- déterminer le score que l'apprenant devra obtenir au [nom de l'évaluation] pour démontrer qu'il a atteint les attentes mondiales pour [niveaux scolaires]
- permettre la communication des résultats de l'évaluation [nom de l'évaluation] au niveau international

Slide 8

DOSSIER DU PARTICIPANT 

1. Programme
2. Identifiant du panéliste
3. Glossaire des termes
4. Liste des acronymes
5. Descripteurs pour le [niveau/matière pertinent] du GFP
6. Instrument(s) d'évaluation [nom de l'évaluation]
7. Diapositives (imprimées sous format notes)
8. Formulaire d'évaluation d'alignement
9. Formulaire d'évaluation des items

PRÉSENTATION 2 – APERÇU DE L'AGENDA, DES OBJECTIFS ET DE LA VUE D'ENSEMBLE METHODOLOGIQUE

Slide 9

OBJECTIFS CLÉS DE CETTE SESSION

- Partager l'agenda de l'atelier
- Comprendre le but de l'établissement de liens entre les politiques

Slide 10

APERÇU DE L'ATELIER

Jour 1-[Date]	Jour 4-[Date]
Ouverture, présentations, logistique et agenda	Présentation de la tâche 2 : Correspondance des résultats
Contexte, objectif et tâches	Présentation de la tâche 3 : Évaluation mondiale et méthode Angoff
Présentation générale : Établissement de liens entre les politiques et le GPF	Activité de la tâche 3 : Pratique de l'évaluation Angoff
Jour 2-[Date]	Jour 5-[Date]
Présentation générale : [nom de l'évaluation]	Activité de la tâche 3 : Réalisation de la série 1 Angoff
Présentation de la tâche 1 : GPF et alignement	Présentation de la tâche 3 : Résultats de la série 1
Activité de la tâche 1 : Aligner l'évaluation avec le GPF	Présentation de la tâche 3 : Méthode Angoff (révision)
Jour 3-[Date]	Jour 6-[Date]
Présentation de la tâche 1 : Résultats d'alignement	Activité de la tâche 3 : Conclure la série 2 de la méthode Angoff
Présentation de la tâche 2 : Mise en correspondance des évaluations et du descripteur mondial de compétences/Niveaux (GPDGPR)	Activité de la tâche 3 : Évaluation de l'atelier
Activité de la tâche 2 : Faire correspondre l'évaluation avec les GPDGPR	Présentation de la tâche 3 : Résultats de la série 2
	Closure et logistique

Slide 11

PRÉSENTATION

QU'EST-CE QUE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES ?

Slide 12

OBJECTIF DE DEVELOPPEMENT DURABLE 4.1.1

En 2015, les Nations Unies ont adopté 17 objectifs de développement durable (ODD) – l'ODD 4 concerne l'éducation de qualité.

La cible 4.1 vise à garantir que toutes les filles et tous les garçons suivent une éducation primaire et secondaire gratuite, équitable et de qualité conduisant à des résultats d'apprentissage pertinents et efficaces d'ici à 2030

Pour mesurer la réussite, l'ODD 4.1.1 vise à quantifier la « Proportion d'enfants et de jeunes : (a) en 2ème/3ème année ; (b) à la fin du primaire ; et (c) à la fin du secondaire inférieur atteignant au moins un niveau de compétence minimum en (i) lecture et (ii) mathématiques, par sexe ».

Slide 13

QU'EST-CE QUE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES ?

L'établissement de liens entre les politiques est le processus que nous utiliserons pour déterminer le niveau de compétence minimum requis pour [nom de l'évaluation], en se basant sur l'expertise des spécialistes (c'est-à-dire vous !)

Pour ce faire, nous allons :

- Vérifier dans quelle mesure le [nom de l'évaluation] est en accord avec les attentes mondiales en [matières] pour [niveaux scolaires], en utilisant un document appelé le Cadre Mondial de Compétences (GPF)
- Fixer le score que l'apprenant doit obtenir au [nom de l'évaluation] pour démontrer qu'il a atteint le niveau de compétence minimum requis pour [niveaux scolaires]

Slide 14

TERMINOLOGIE CLÉ

- **Cadre mondial de compétences (GPF)** = un document qui décrit les attentes mondialement acceptées à la fin de chaque niveau en lecture et en mathématiques
- **Niveaux de compétence minimum (Minimum Proficiency Levels, MPL)** = les normes minimales attendues des apprenants à la fin des 2ème/3ème années, à la fin du primaire et à la fin du secondaire inférieur
- **Alignement** = si le contenu d'une évaluation est similaire aux définitions mondialement acceptées de ce qui constitue la lecture et les mathématiques
- **Repère** = le score qu'un apprenant doit obtenir à une évaluation pour démontrer qu'il atteint un standard particulier, par exemple, le MPL
- **Descripteur mondial de compétences (GPD)** = Une définition détaillée élaborée par des experts en la matière qui explique quoi quelle mesure le contenu décrit dans les domaines de connaissance et de compétence du GPF un apprenant devrait être capable de démontrer dans une matière à un niveau scolaire.

Slide 15

CALENDRIER DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

- **Septembre 2017** : Un atelier des parties prenantes de l'Institut de statistique de l'UNESCO (ISU) a proposé l'établissement de liens entre les politiques comme méthode pour établir des repères mondiaux pour chaque évaluation, basé sur une échelle commune de compétence
- **Août 2018** : un atelier conjoint de l'Agence des États-Unis pour le développement international (USAID) et de l'ISU a discuté de l'établissement de liens entre les politiques pour la communication des compétences minimales grâce à l'ODD 4.1.1 et les indicateurs de l'USAID
- **Avril/Mai 2019** : Cadre Mondial de Compétences (GPF) rédigé
- **Septembre 2019** : Rédaction du Projet de Boîte à Outils d'Établissement de Liens entre les Politiques pour Mesurer les Résultats Mondiaux d'Apprentissage (PLT)

Slide 16

ANTÉCÉDENTS DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES : CALENDRIER

- **Octobre 2019 – Septembre 2020** : Cinq ateliers pilotes réalisés
- **Juin – Octobre 2020** : Mise à jour du GPF et du PLT basée sur les ateliers pilotes
- **Octobre 2020 - Août 2022** : Organisation d'ateliers supplémentaires pour poursuivre la phase pilote du GPF et du PLT
- **Août 2022 - Mars 2023** : Mise à jour du PLT

Slide 17

LES TÂCHES CLÉS POUR L'ATELIER D'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

- Familiarisation** : Se familiariser avec le GPF et le [nom de l'évaluation]
- Alignement et Correspondance** : Vérifier l'alignement entre le [nom de l'évaluation] et le GPF (TÂCHE 1) et faire correspondre les items de l'évaluation avec les descripteurs du GPF (TÂCHE 2)
- Référencement** : Établir les repères sur le [nom de l'évaluation] grâce à deux séries d'évaluation (TÂCHE 3)

Slide 18

ÉTABLISSEMENT DE RÉFÉRENCES MONDIALES POUR PLUSIEURS ÉVALUATIONS

- L'établissement de **références mondiales** pour différentes évaluations relie chaque évaluation au GPF.
- Le positionnement des repères mondiaux sur l'échelle d'évaluation dépend de la difficulté de l'évaluation par rapport au GPF, comme déterminé par les jugements des panelistes.

Process flow: Déterminer les paramètres de la distribution mondiale standardisée → Paramétriser l'échelle à l'aide de la distribution mondiale standardisée → Définir la compétence mondiale standard → Définir la compétence mondiale standard

Slide 19

AVANTAGES DE L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES

- Permettre **trois types d'analyses** avec les repères mondiaux :
 - **Comparer** les résultats des évaluations à travers différents contextes/langues à l'intérieur du pays et avec les résultats d'autres pays
 - **Agréger** les résultats des évaluations pour différentes évaluations dans le pays et avec ceux d'autres pays
 - **Suivre les résultats des évaluations dans le temps pour surveiller les progrès**
- Permettre la prise en charge nationale des résultats - des repères fixés par les pays pour les pays.
- Déterminer si les apprenants ont développé les connaissances et les compétences attendues pour leur niveau scolaire.

Présentation 3 – Familiarisation avec le Gpf

Slide 20

PRÉSENTATION

QU'EST-CE QUE LE CADRE MONDIAL DE COMPÉTENCES (GPF) ?

Slide 21

LE CADRE MONDIAL DE COMPÉTENCES

- Créé par des experts mondiaux en lecture et en mathématiques, et révisé en fonction des ateliers pilotes
- Établit la compétence minimale mondiale (ce que les apprenants doivent être en mesure de connaître et de faire) en lecture et en mathématiques pour les 1ère à 9e année
- Fondé sur des données probantes et s'appuyant sur :
 - des progressions développementales
 - des données issues des programmes scolaires et des cadres d'évaluation d'une cinquantaine de pays
- Non prescriptif

Slide 22

NIVEAUX MONDIAUX DE COMPÉTENCES (GPL)

Dans le cadre de leur travail de communication des résultats en lien avec l'Objectif de Développement Durable 4.1.1, l'UNESCO-ISU et ses partenaires ont établi quatre Niveaux Mondiaux de Compétences (GPL) pour le GPF :

- En dessous de partiellement atteint la compétence minimale mondiale
- Partiellement atteint la compétence minimale mondiale
- Atteint la compétence minimale mondiale
- Dépasse la compétence minimale mondiale

Slide 23

NIVEAUX MONDIAUX DE COMPÉTENCES (GPL)

Dans le cadre de leur travail de communication des résultats en lien avec l'Objectif de Développement Durable 4.1.1, l'UNESCO-ISU et ses partenaires ont établi quatre Niveaux Mondiaux de Compétences (GPL) :

- En dessous de partiellement atteint la compétence minimale mondiale
- Partiellement atteint la compétence minimale mondiale
- **Atteint la compétence minimale mondiale** ← GPL utilisé pour la communication des résultats de l'ODD 4.1.1
- Dépasse la compétence minimale mondiale

Slide 39

ACTIVITÉ D'ÉVALUATION 

- Comment l'activité d'évaluation préalable à l'atelier s'est-elle déroulée ?
- Avez-vous pu évaluer :
 - 3 apprenants que vous avez classés comme ayant atteint partiellement le niveau mondial de compétences minimal
 - 3 apprenants qui atteignent le niveau mondial de compétences minimal
 - 3 apprenants qui dépassent le niveau de compétence minimal mondial ?
- Quels sont les résultats des apprenants à l'évaluation ?
 - Sur quels items ont-ils bien réussi et quels items ont été plus difficiles ?
 - Quelles étaient quelques-unes des erreurs typiques qu'ils ont commises ?

Présentation 5 – Former les membres de l'équipe sur la tâche d'alignement

Slide 40

PRÉSENTATION 

TÂCHE 1 : VÉRIFIER L'ALIGNEMENT DU CONTENU ENTRE [NOM DE L'ÉVALUATION] ET LE GPF

Slide 41

L'ÉTUDE D'ALIGNEMENT 

- **Activité 1**—La première activité de l'atelier.
- **Tâche**—Les membres de l'équipe feront des évaluations individuelles et indépendantes pour déterminer si les items de [nom de l'évaluation] sont alignés avec le [niveau pertinent] du GPF.
- **Objectif**—S'assurer que les membres de l'équipe ont bien compris le GPF et leur permettre d'identifier les énoncés de connaissances et/ou de compétences décrivant les connaissances et/ou compétences requises des enfants pour répondre correctement aux items de l'évaluation.
- **Alignement suffisant**—L'alignement est important pour s'assurer qu'il y a suffisamment d'items dans une évaluation qui mesurent les connaissances et/ou compétences dépeintes dans le GPF afin que l'établissement de liens entre les politiques fonctionne.

Slide 42

ALIGNEMENT DE [NOM DE L'ÉVALUATION] ET DU GPF 

Il y a **deux étapes principales**—chacune avec des sous-étapes—pour l'alignement.

- **Étape 1** : Les membres de l'équipe évaluent indépendamment l'alignement entre les items de [nom de l'évaluation] et les énoncés de connaissances et/ou de compétences du GPF en utilisant un processus en trois sous-étapes.
- **Étape 2** : Les animateurs rassemblent et résument les évaluations pour vérifier l'alignement entre les évaluations et le GPF.

Slide 43

ALIGNEMENT DE [NOM DE L'ÉVALUATION] ET DU GPF 

Étape 1 (complétée par les membres de l'équipe)

- Pratiquez la réalisation d'évaluations des énoncés de connaissances et/ou de compétences avec des exemples d'items.
- Travaillez individuellement et indépendamment pour évaluer l'alignement entre chaque item d'évaluation et les énoncés de connaissances et/ou de compétences du GPF.
- Commencez par le premier item et procédez item par item ; trouvez les énoncés de connaissances et/ou de compétences du GPF qui s'alignent (le cas échéant) avec les connaissances ou compétences nécessaires pour répondre correctement à l'item.
- Enregistrez les évaluations sur le formulaire d'évaluation d'alignement en utilisant l'échelle d'évaluation (sur la diapositive suivante).

Slide 44



ALIGNEMENT DE [NOM DE L'ÉVALUATION] ET DU GPF

Évaluez chaque item en utilisant une échelle de **Correspondance complète**, **Correspondance partielle** et **Sans correspondance** comme suit :

- **Correspondance complète (C)** signifie que tout le contenu nécessaire pour répondre correctement à l'item est contenu dans l'énoncé de connaissances et/ou de compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il utilise complètement les connaissances et/ou compétences décrites dans l'énoncé.
- **Correspondance partielle (P)** signifie qu'une partie du contenu nécessaire pour répondre correctement à l'item est contenue dans l'énoncé de connaissances et/ou de compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il utilise partiellement les connaissances et/ou compétences décrites dans l'énoncé.
- **Sans correspondance (N)** signifie qu'aucune partie du contenu nécessaire pour répondre correctement à l'item n'est contenue dans les énoncés de connaissances et/ou de compétences, c'est-à-dire que si l'apprenant répond correctement à l'item, c'est parce qu'il n'utilise pas les connaissances et/ou compétences décrites dans le GPF.

Slide 45



ALIGNEMENT DE [NOM DE L'ÉVALUATION] ET DU GPF

Suivez ces **instructions supplémentaires** pour les évaluations d'alignement :

- Si un item a une évaluation de **correspondance complète (C)** avec un énoncé spécifique de connaissances et/ou de compétences, les membres de l'équipe ne doivent pas l'associer à d'autres énoncés de connaissance et/ou de compétence, ce qui signifie qu'il est aligné uniquement avec un énoncé dans le GPF ;
- Si un item a une évaluation de **correspondance partielle (P)** avec un énoncé spécifique de connaissances et/ou de compétences, les membres de l'équipe devraient généralement l'associer à un ou deux énoncés supplémentaires de connaissances et/ou de compétences ; et
- Si un item a une évaluation **sans correspondance (N)** avec n'importe quel énoncé de connaissances et/ou de compétences, les membres de l'équipe ne devraient l'associer à aucun énoncé de connaissances et/ou de compétences.

Slide 46



EXEMPLE : CORRESPONDANCE COMPLÈTE

<p>I. Comment s'écrit huit cent soixante-dix sous forme normalisée ?</p> <p>A. 807 B. 870 C. 817 D. 871</p>	<p>Domaine : Nombres et opérations Concept : Nombres Entiers Sous-concept : Identifier, compter et déterminer l'importance relative des nombres entiers Énoncé de connaissances ou de compétences : Compter, lire et écrire des nombres entiers</p>
---	--

Pour répondre correctement à cet item, l'apprenant doit être capable d'identifier et de compter des nombres entiers. Par conséquent, l'item peut être évalué comme une « correspondance complète » avec l'énoncé de connaissances et/ou de compétences, car il nécessite uniquement les connaissances ou les compétences de cet énoncé spécifique.

Slide 47



EXEMPLE : CORRESPONDANCE PARTIELLE

<p>2. Quelle est la somme la plus grande ?</p> <p>A. 22 + 37 B. 21 + 39 C. 23 + 38 D. 24 + 36</p>	<p>Domaine : Nombres et Opérations Concept : Nombres Entiers Sous-concept : Résoudre des opérations en utilisant des nombres entiers Énoncé de connaissances ou de compétences : Additionner, soustraire, multiplier et diviser des nombres entiers</p>
---	--

Pour répondre correctement à cet item, l'apprenant doit être capable de comparer et d'ordonner des nombres entiers ainsi que d'additionner et de soustraire des nombres entiers. Par conséquent, l'item peut être évalué comme une « correspondance partielle » avec les énoncés de connaissances et/ou de compétences, car il nécessite des connaissances ou des compétences issues des deux énoncés.

<p>Domaine : Nombres et Opérations Concept : Nombres Entiers Sous-concept : Identifier, compter et déterminer l'importance relative des nombres entiers Énoncé de connaissances ou de compétences : Comparer et ordonner des nombres entiers</p>

Slide 48



EXEMPLE : SANS CORRESPONDANCE

<p>3. Qu'est-ce que $2/3 - 1/3$?</p> <p>A. 1/0 B. 1/3 C. 2/3 D. 3/0</p>	<p>Pour répondre correctement à cet item, l'apprenant doit être capable d'additionner et de soustraire des fractions. Cette connaissance ou compétence n'est pas attendue avant les niveaux d'enseignement primaire supérieur. Par conséquent, l'item peut être évalué comme « sans correspondance », car il requiert une connaissance ou compétence qui n'est pas attendue au niveau scolaire (ou avant).</p>
---	--

Slide 54

NIVEAUX D'ALIGNEMENT EN TERMES D'ÉTENDUE ET DE PROFONDEUR

Une fois la tâche d'alignement terminée, nous examinerons l'étendue et la profondeur de l'alignement par rapport au GPF :

- La **profondeur** concerne le nombre de domaines couverts dans l'évaluation
- L'**étendue** concerne le nombre de sous-concepts du GPF couvertes dans l'évaluation

L'évaluation doit être convenablement alignée pour produire des rapports conformément à l'ODD 4.1.1

Slide 55

ALIGNEMENT DE [NOM DE L'EVALUATION] ET DU GPF 

Étape 2 (complétée par les animateurs)

- Compiler, analyser et résumer les évaluations d'alignement
- Calculer les totaux, les moyennes et les médianes
- Déterminer si l'évaluation est convenablement alignée en termes d'étendue et de profondeur

PRÉSENTATION 6 – LES MEMBRES DE L'ÉQUIPE EFFECTUENT LA TACHE D'ALIGNEMENT

Slide 56

ACTIVITÉ 

LES MEMBRES DE L'ÉQUIPE S'ALIGNENT SUR [NOM DE L'EVALUATION] ET LE GPF

Slide 57

ALIGNEMENT DE [NOM DE L'EVALUATION] ET DU GPF 

Étape 1 (complétée par les membres de l'équipe)

- Pratiquer la réalisation des évaluations d'alignement avec des exemples d'items
- Travailler indépendamment pour évaluer l'alignement entre chaque item de l'UE et les énoncés du GPF relatifs aux connaissances et/ou compétences
- Commencer par le premier item et procéder item par item ; trouver les énoncés du GPF relatifs aux connaissances et/ou compétences qui s'alignent (le cas échéant) avec les connaissances ou compétences nécessaires pour répondre correctement à l'item.
- Enregistrer les évaluations sur le formulaire d'évaluation d'alignement en utilisant l'échelle d'évaluation :
 - Correspondance complète
 - Correspondance partielle
 - Sans correspondance

PRÉSENTATION 7 – PRÉSENTATION ET DISCUSSION DES RÉSULTATS DE LA TÂCHE D'ALIGNEMENT

Slide 58

PRÉSENTATION

RÉSULTATS DE L'ALIGNEMENT DES MEMBRES DES ÉQUIPES DE LA TÂCHE I

Slide 59

NOTES D'ALIGNEMENT

Domaine	Items
N	14
M	7
G	3
S	2
A	0
Total	26
Concept	Items
N1	14
M1	3
M2	4
M3	0
OT	2
O2	0
O3	1
ST	2
A1	0
A3	0
Total	26

Slide 60

NOTES D'ALIGNEMENT

Sous-concept	Items
N1.1	4
N1.2	0
N1.3	8
N1.4	2
M1.1	3
M2.1	2
M2.2	2
M3.1	0
G1.1	2
G2.1	0
G3.1	1
S1.1	2
A1.1	0
A3.2	0
Total	26

Slide 61

NOTES D'ALIGNEMENT

Domaine	Items
C	14
D	7
R	5
Total	24
Concept	Items
C1	14
C2	0
C3	3
D1	4
D2	0
R1	2
R2	3
R3	1
Total	24

Slide 62

NOTES D'ALIGNEMENT

Sous-concept	Items
C1.1	14
C1.2	0
C2.1	0
C3.1	2
D1.1	0
D1.2	3
D2.1	2
R1.1	2
R1.2	0
R1.3	2
R2.1	0
R2.2	1
R2.3	2
R3.1	0
R3.2	0
R3.3	0
Total	26

Slide 63

NOTES D'ALIGNEMENT

Certaines personnes ont aligné cet item sur :

- **Domaine** : Mesure
- **Concept** : Longueur, poids, capacité, volume, surface et périmètre
- **Sous-concept** : Utiliser des unités non standard et standard pour mesurer, comparer et ordonner
- **Connaissances et/ou compétences** : Utiliser des unités non standard pour estimer, mesurer et comparer la longueur, le poids, le volume et la capacité
- **Correspondance** : Complète

D'autres l'ont aligné sur :

- **Domaine** : Géométrie
- **Concept** : Position et direction
- **Sous-concept** : Décrire la position et la direction des objets dans l'espace
- **Connaissances et/ou compétences** : Utiliser des termes de position pour décrire l'emplacement d'un objet
- **Correspondance** : Complète

PRÉSENTATION 8 – FORMER LES MEMBRES DE L'ÉQUIPE A LA TACHE DE MISE EN CORRESPONDANCE

Slide 64

PRÉSENTATION

TÂCHE 2. ASSOCIER LES ITEMS DE [NOM DE L'ÉVALUATION] AVEC LES NIVEAUX DE COMPÉTENCE ET LES DESCRIPTEURS DANS LE GPF

Slide 65

CADRE MONDIAL DE COMPÉTENCES (RÉVISION)

Le GPF dispose de **GPL** (niveaux) et de **GPD** (descripteurs) pour les niveaux 1 à 9 en lecture et en mathématiques :

- Les GPD décrivent le **niveau de compétence minimal** requis pour les GPL, c'est-à-dire les connaissances ou compétences minimales nécessaires pour être classé dans chaque GPL (par niveau et matière).
- Les GPD sont organisés hiérarchiquement par **domaines**, **concepts**, **sous-concepts**, ainsi que **connaissances et compétences**, avec des descripteurs pour chacune des connaissances et compétences.

Slide 66

CADRE MONDIAL DE COMPÉTENCES (EXEMPLE)

Partially Meets Global Minimum Proficiency	Meets Global Minimum Proficiency	Exceeds Global Minimum Proficiency
M2 : HEURE M2.1.1_P Identifier, classer et décrire les activités/livements qui ont lieu à différents moments de la journée (par exemple, le matin et l'après-midi).	M2.1.1_M Indiquer l'heure en utilisant une horloge analogique à l'heure la plus proche.	M2.1.1_E Indiquer l'heure en utilisant une horloge analogique à la demi-heure la plus proche.
M2.1.2_P ND	M2.1.2_M Reconnaître le nombre de jours dans une semaine et de mois dans une année.	M2.1.2_E Reconnaître le nombre d'heures dans une journée, le nombre de minutes dans une heure et le nombre de secondes dans une minute.

Slide 67

CADRE MONDIAL DE COMPÉTENCES (EXEMPLE)

Partiellement atteint la compétence minimale mondiale	Atteint le niveau mondial de compétences minimal	Dépasse le niveau mondial de compétences minimal
R. COMPRÉHENSION DE LA LECTURE		
RT : RECUPERER DES INFORMATIONS		
RT.1.1_P Reconnaître le sens des mots les plus fréquents de niveau 2 (e.g., match à given word for an illustration (or synonym or provide a brief spoken/written definition).	RT.1.1_M Reconnaître le sens des mots fréquents de niveau 2. (Par exemple, associer un mot donné à une illustration ou à un synonyme, ou fournir une brève définition parlée/écrite).	RT.1.1_E Reconnaître le sens des mots moins fréquents de niveau 2. (Par exemple, associer un mot donné à une illustration ou à un synonyme, ou fournir une brève définition parlée/écrite).

Slide 68

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD 

- **Approfondissez** votre compréhension des items de [nom de l'évaluation] et du GPF acquis lors de l'activité d'alignement dans la Tâche 1.
- **Activité en groupe** : Vous devriez travailler pour parvenir à un consensus.
- **Concentrez-vous sur un aspect clé** : Descripteurs (GPD) mondiaux de compétences minimum qui correspondent aux items.

Slide 69

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

Répondez à ces questions pour chaque item (en fonction du consensus dans les groupes) :

- Quelles **connaissances et/ou compétences** sont nécessaires pour répondre correctement aux items ?
- Qu'est-ce qui rend l'item **facile ou difficile** ?
- Quels sont les plus **faibles GPL** et GPD qui sont les plus appropriés pour l'item ?

Slide 70

ASSOCIATION DES ITEMS « sans correspondance »

Pour les items « sans correspondance », nous suivrons ces étapes :

- Imaginez un groupe d'apprenants qui sont le mieux décrits par les GPD de niveau « partiellement atteint ».
- En utilisant votre expérience d'enseignement à de tels apprenants, déterminez si cet item convient à ces apprenants et s'ils seraient susceptibles de répondre correctement à l'item.
- Si l'item est jugé approprié pour les apprenants au niveau « partiellement atteint », cela peut être enregistré.
- Sinon, le processus devrait être répété pour le niveau « atteint » et ensuite pour le niveau « dépasse », si nécessaire.
- Si l'item est jugé très difficile pour le niveau, alors il devrait être enregistré comme étant au-dessus du niveau « dépasse ».

Slide 71

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD 

L'item provient d'une évaluation de niveau 3 et est donc associé au niveau 2 pour le rapport sur l'ODD 4.1.1(a). Il correspond au niveau « Atteint » des GPL et GPD (norme de performance) de niveau 2.

<p>Comment écrit-on quatre-vingt-sept sous forme normalisée ?</p> <p>A. 90 B. 87 C. 807 D. 870</p> <p>Domaine : Nombres et opérations Concept : Nombres Entiers Sous-concept : Identifier et compter les nombres entiers, et identifier leur importance relative Connaissances ou compétences (norme de contenu) : Compter, lire et écrire des nombres entiers</p>	<p>Qu'est-ce qui le rend facile ou difficile : les autres choix de réponses sont des distractions majeures, en particulier C.</p> <p>GPL et GPD (norme de performance) :</p> <p>Partiellement atteint : Lire et écrire les nombres entiers jusqu'à 30 en lettres et en chiffres.</p> <p>Atteint : Lire et écrire les nombres entiers jusqu'à 100 en lettres et en chiffres.</p> <p>Dépasse : N/D</p>
---	---

Slide 72

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD 

Job avait 16 pêches.
Il en a donné 4.
Ensuite, Job a divisé les pêches restantes en parts égales dans 2 paniers.
Combien de pêches Job a-t-il mis dans chaque panier ?

A. 6
B. 8
C. 10
D. 12

Source: Malin, V. S., Marry, M. G., Fay, P., & Hooper, M. (2010). *IM2012*. www.illustrativemathematics.org

<p>Domaine : Nombres et Opérations Concept : Nombres Entiers Sous-concept : Résoudre des problèmes du monde réel impliquant des nombres entiers Connaissances/compétences : Résoudre des problèmes du monde réel impliquant l'addition, la soustraction, la multiplication et la division de nombres entiers.</p>	<p>Qu'est-ce qui rend l'item facile/difficile ?</p> <p>Difficile—Comme il s'agit d'un problème réel, les apprenants doivent identifier quelles opérations doivent être effectuées ; il y a deux opérations/étapes à réaliser.</p> <p>Quel est le GPD le plus bas pour répondre correctement ?</p> <p>Au-dessus du niveau « Dépasse » pour le niveau 3 car il correspond au niveau « Atteint » du niveau 4—Résoudre des problèmes simples du monde réel impliquant la multiplication de deux nombres entiers jusqu'à 5, ainsi que</p>
---	---

Slide 73

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

Quelle est la différence d'heure affichée entre ces deux horloges ?



Qu'est-ce qui rend l'item facile/difficile ?
Difficile—C'est un problème en deux étapes et les nombres ne sont pas indiqués sur les horloges.

Quel est le GPD le plus bas pour répondre correctement ?
Bien que « Indiquer l'heure à l'aide d'une horloge analogique à l'heure la plus proche » soit considéré comme « atteint » au niveau 2, les élèves du niveau 2 ne sont pas censés résoudre des problèmes impliquant le temps ; c'est ce qui est attendu au niveau « atteint » du niveau 3.

Domaine : Mesure
Concept : Heure
Sous-concept : Indiquer l'heure ET résoudre des problèmes impliquant l'heure
Connaissances ou compétences (norme de contenu) : Indiquer l'heure à l'aide d'une horloge analogique ET identifier ou résoudre des problèmes impliquant des équivalences entre différentes unités de temps

Slide 74

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

Quel rectangle est ombré à $\frac{1}{3}$?



Qu'est-ce qui rend l'item facile ou difficile ?
Difficile—comprendre que la portion ombrée doit être de 1/3 ombrée plutôt que simplement 1 des 3 parties doit être ombrée. C'est une distraction majeure.

Quel est le GPD le plus bas pour répondre correctement ?
Bien que cet item corresponde partiellement au niveau 3—(Identifie les fractions unitaires quotidiennes représentées sous forme d'objets ou d'images en notation fractionnaire), il est « sans correspondance » au niveau 2

Domaine : Nombres et Opérations
Concept : Fractions
Sous-concept : Identifier et représenter des fractions à l'aide d'objets, d'images et de symboles, et identifier leur importance relative
Connaissances/compétences : Exprimer une représentation visuelle d'une fraction (image, objets) en notation fractionnaire

Slide 75

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

Cet item utilise le texte du diaporama 49 de la présentation 5. Le texte est adapté au niveau 3, il faut donc en tenir compte lors de la détermination de la correspondance.

Item : Qui possède un chien de compagnie ?

GPL et GPD (norme de performance) :
Partiellement atteint : Récupérer une seule information importante et explicite à partir d'un texte de niveau 3 en effectuant une correspondance directe ou proche de mots lorsque l'information requise est adjacente au mot correspondant et qu'il n'y a pas d'informations contraires.
Atteint : « ... » et il y a peu d'informations contraires « ... »
Dépasse : Récupérer plusieurs éléments d'information explicite ... lorsque l'information requise est adjacente au mot correspondant et qu'il y a peu d'informations contraires ...

Domaine : Compréhension de la lecture
Concept : Récupérer des informations
Sous-concept : Récupérer des informations explicites dans un texte de niveau scolaire en utilisant une correspondance directe ou proche de mots
Connaissances ou compétences : Récupérer une seule information explicite à partir d'un texte continu de niveau scolaire en utilisant une correspondance directe ou proche

Slide 76

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

L'item correspond à cette déclaration de connaissances et/ou de compétences de niveau 3 et au niveau de performance « partiellement atteint » des GPL et des GPD (norme de performance).

Item : Passage de décodage — Mot « Jabu »

GPL et GPD (norme de performance) :
Partiellement atteint : Lire ou signer avec précision un texte continu de niveau 3, à un rythme considéré lent selon les normes du pays pour la fluidité dans la langue dans laquelle l'évaluation est administrée (par exemple, souvent mot par mot).
Atteint : Lire ou signer avec précision un texte continu de niveau 3, à un rythme conforme aux normes minimales du pays pour la fluidité dans la langue dans laquelle l'évaluation est administrée.
Dépasse : Lire ou signer avec précision un texte continu de niveau 3, à un rythme supérieur aux normes minimales du pays pour la fluidité dans la langue dans laquelle l'évaluation est administrée.

Domaine : Décodage
Concept : Fluidité
Sous-concept : Lire ou signer un texte continu de niveau scolaire à un rythme et avec précision
Connaissances ou compétences : Lire ou signer avec fluidité un texte continu de niveau scolaire

Slide 77

ASSOCIATION D'ITEMS AVEC LES GPL ET GPD

L'item correspond à cette déclaration de connaissances et/ou compétences de niveau 3 et au niveau « dépasse » des GPL et des GPD (norme de performance).

Item : Pourquoi le chien est-il revenu ?

GPL et GPD (norme de performance) :
Partiellement atteint : Faire des inférences simples dans un texte de niveau 3 en reliant deux informations explicites dans des phrases consécutives lorsqu'il n'y a pas d'information contraire. Cela sera généralement le cas en réponse à une question « pourquoi » ou « comment ». (Voir les exemples d'articles à l'annexe C).
Atteint : « ... » lorsque les informations contraires sont limitées « ... »
Dépasse : « ... » dans un ou plusieurs paragraphes lorsque les informations à relier sont plus éloignées et/ou qu'il y a beaucoup d'informations contraires « ... »

Domaine : Compréhension de la lecture
Concept : Interpréter l'information
Sous-concept : Faire des déductions dans un texte de niveau scolaire
Connaissances ou compétences : Faire des déductions simples dans un texte de niveau scolaire en reliant des informations explicites et/ou implicites dans le texte

PRÉSENTATION 9 – LES MEMBRES DE L'EQUIPE REALISENT LA TACHE DE MISE EN CORRESPONDANCE

Slide 78

ACTIVITÉ 

LES MEMBRES DES ÉQUIPES METTENT EN CORRESPONDANCE LES ITEMS DE [NOM DE L'ÉVALUATION] AVEC LES NIVEAUX ET LES DESCRIPTEURS DU GPF

Slide 79

MISE EN CORRESPONDANCE DES ITEMS AVEC LES GPL ET GPD

1. **Travailler en groupes au niveau des membres de l'équipe** : commencer par le premier item de l'évaluation et procéder item par item.
2. **Examiner la connaissance ou la compétence dans le GPF** (à partir de la Tâche 1) qui correspond à chaque item.
3. **Parvenir à un consensus sur l'énoncé des connaissances et/ou des compétences requises et le niveau GPL et GPD** (niveau de performance) nécessaires pour répondre correctement au mot, à la question ou à l'item.
4. **Identifier également ce qui rend l'item facile ou difficile.**
5. **Écrire le niveau GPL, le descripteur GPD et ce qui rend l'item facile ou difficile** sur le livret de test à côté du numéro de l'item, de la question ou du mot sur le GPF qui correspond à l'item.

PRÉSENTATION 10 – PRÉSENTATION ET DISCUSSION DES RESULTATS DE LA MISE EN CORRESPONDANCE DES TACHES

Slide 80

PRÉSENTATION

EXAMEN DES MEMBRES DE L'ÉQUIPE
CORRESPONDANCE DES
RÉSULTATS DE LA TÂCHE 2

Slide 81

DISCUSSION SUR LA TÂCHE 2 DE CORRESPONDANCE DES GROUPES

1. **Vous êtes-vous concentré sur cet aspect clé ?**
Correspondance des descriptions des niveaux des apprenants ayant un niveau de compétence minimal mondial (GPL et GPD) avec les items
2. **Avez-vous rencontré des difficultés à parvenir à un consensus sur certains items ? Si oui, lesquels et pourquoi ?**
3. **Étiez-vous tous d'accord avec les décisions du groupe ? Pourquoi ou pourquoi pas ?**

Slide 82

RÉSULTATS DE LA CORRESPONDANCE DES GROUPES POUR LA TÂCHE 2

Passons en revue des résultats de la correspondance de votre groupe sur les items de [nom de l'évaluation]

Résumer les réponses à ces questions pour chaque item (en fonction du consensus du groupe) :

1. Quelles **connaissances et compétences** sont nécessaires pour répondre correctement aux items ?
2. L'item est-il **facile ou difficile** ?
3. Quel est le niveau **GPL et GPD** le plus bas qui est le plus approprié pour l'item ?

PRÉSENTATION 11 – APERÇU DE L'APPROCHE DES NORMES MONDIALES ET DES RÉFÉRENCES

Slide 83

PRÉSENTATION

TÂCHE 3. ÉTABLISSEZ DES RÉFÉRENCES MONDIALES POUR [NOM DE L'ÉVALUATION]

Slide 84

ÉTABLISSEMENT DE RÉFÉRENCES MONDIALES

- Utiliser une procédure d'évaluation standardisée (la **méthode Angoff modifiée**) pour établir des référentiels mondiaux qui relient l'évaluation [nom de l'évaluation] au GPF.
- Concentrez-vous sur l'établissement de la **référence « Atteint »** pour séparer les scores de l'évaluation [nom de l'évaluation] en deux niveaux.
- Par exemple, imaginer une référence « Atteint » de 50 points sur une **échelle** de 0 à 100 points.
- Déterminer les **plages de notes pour les deux niveaux** :
 - **En dessous de Partiellement atteint** / Partiellement atteint = de 0 à 49 points
 - **Atteint/Dépasse** = de 50 à 100 points.

Slide 85

ÉTABLISSEMENT DE RÉFÉRENCES MONDIALES POUR PLUSIEURS ÉVALUATIONS

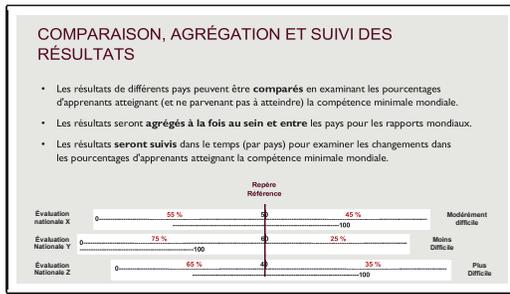
- L'établissement de **références mondiales** pour différentes évaluations relie chaque évaluation au GPF.
- La positionnement des repères mondiaux sur l'échelle d'évaluation dépend de la **difficulté** de l'évaluation par rapport au GPF, comme déterminé par les jugements des panélistes.

Slide 86

CALCUL DES POURCENTAGES DE COMPÉTENCE MINIMALE MONDIALE

- En appliquant les références mondiales aux données (et en généralisant à partir d'un échantillon) pour chaque évaluation, on obtient les **pourcentages d'apprenants** atteignant les compétences minimales mondiales.
- Il est **nécessaire** de rendre compte de ces pourcentages pour les indicateurs ODD [et USAID].

Slide 87

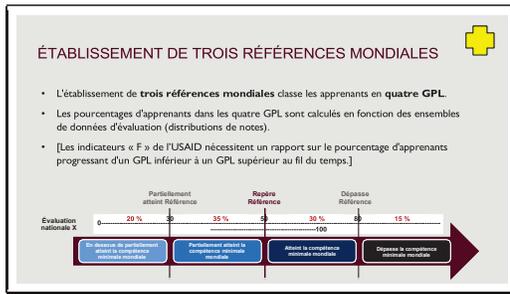


Slide 88

COMPARAISON, AGRÉGATION ET SUIVI DES RÉSULTATS

Pays et Évaluation	Niveaux de Compétence Minimum Mondial			
	En dessous de partiellement atteint/Partiellement atteint		Atteint/Dépasse	
	Plage de notes	Pourcentage	Plage de notes	Pourcentage
Évaluation nationale X	0-49	55 %	50-100	45 %
Évaluation Nationale Y	0-59	75 %	60-100	25 %
Évaluation Nationale Z	0-39	65 %	40-100	35 %

Slide 89



PRÉSENTATION 12 – FORMER LES MEMBRES DE L'EQUIPE SUR LA TACHE D'ÉVALUATION

Slide 90

PRÉSENTATION

MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

Slide 91

MISE EN ŒUVRE DE LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

- La **méthode Angoff modifiée** est utilisée pour définir les références :
 - Méthode d'évaluation largement utilisée
 - Repose sur les jugements des experts des membres de l'équipe
 - Méthode centrée sur les questions, c'est-à-dire que les membres de l'équipe évaluent chaque question en estimant si les apprenants ayant un niveau de compétence minimal à chaque GPL répondraient correctement à la question
- Il est essentiel de se concentrer sur les **définitions de la maîtrise minimale provenant des GPD** dans le GPF au [niveau 2 pour l'ODD 4.1.1(a), niveau 5 pour l'ODD 4.1.1(b) et niveau 8 pour l'ODD 4.1.1(c)]

Slide 92

MISE EN ŒUVRE DE LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

- Les évaluations pour la Tâche 3 doivent être **individuelles et indépendantes**.
- **Le consensus sur les évaluations n'est pas nécessaire**, mais la cohérence est souhaitée.
- **Les références représentent les estimations des membres de l'équipe des notes** qu'un apprenant ayant un niveau de compétence minimal à chaque niveau obtiendrait à l'évaluation.
- Angoff utilise **deux séries** d'évaluations des questions, avec des discussions et des retours entre les cycles.
- **Les références globales** sont calculées en fonction des évaluations totales de chaque membre de l'équipe de travail et des moyennes de l'ensemble des membres de l'équipe de travail.

Slide 93

MISE EN ŒUVRE DE LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

Deux séries

- **Série 1** : Effectuer les **premières évaluations** pour chaque question de l'évaluation.
 - Après la série 1, additionner les évaluations pour calculer les **références mondiales initiales** de chaque membre de l'équipe de travail, puis calculez leur moyenne pour déterminer les références initiales de l'équipe de travail.
- **Série 2** : Effectuer les **évaluations révisées** pour chaque question de l'évaluation.
 - Après la Série 2, additionner les évaluations pour calculer les **références mondiales finales** de chaque membre de l'équipe de travail, puis calculez leur moyenne pour déterminer les références finales de l'équipe de travail.

Slide 94

MISE EN ŒUVRE DE LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

Entre « en dessous de partiellement atteint la compétence minimale mondiale » et « partiellement atteint » la compétence minimale mondiale (partiellement atteint la référence) ➔ Au niveau ou légèrement supérieur au niveau partiellement atteint de la compétence minimale mondiale (juste partiellement atteint ou JPA)

Entre « partiellement atteint » et « atteint » la compétence minimale mondiale (partiellement atteint la référence) ➔ Au niveau ou légèrement au-dessus du niveau de la compétence minimale mondiale (juste atteint ou JA)

Entre « atteint » et « dépasse » la compétence minimale mondiale (dépasse la référence) ➔ Atteint ou légèrement au-dessus du niveau de compétence minimale mondiale (juste dépasse ou JD)

Slide 95

MISE EN ŒUVRE DE LA MÉTHODE ANGOFF POUR L'ÉTABLISSEMENT DES RÉFÉRENCES

- Les évaluations des items sont basées sur **quatre attentes**, c'est-à-dire la probabilité qu'un apprenant ayant le niveau de compétence minimal mondial (basé sur les GDP du GFP) réponde correctement à chaque item :
 - Probablement pas (« non »)
 - Éventuellement (« non »)
 - **Raisonnement sûr ou ≥ 67 % de chance (« oui »)**
 - Absolument sûr (« oui »)
- Les évaluations des items ne sont pas basées sur « devrait » mais sur « serait » pour des **attentes réalistes** :
 - « **Deviat** » se réfère à la performance basée uniquement sur les déclarations de connaissances et/ou de compétences du GFP.
 - « **Serait** » est influencé par les contraintes de l'évaluation, par exemple la difficulté d'un item pour un apprenant particulier, les conditions de passation de l'évaluation, l'anxiété de l'apprenant et les erreurs aléatoires.

Slide 96

SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Juste partiellement atteint (JPA)

Juste atteint (JA)

Juste dépasse (JD)

Étape 1 : Identifier et/ou **conceptualiser** trois apprenants qui ont juste partiellement atteint (JPA) le seuil, trois apprenants qui ont juste atteint (JA) le seuil et trois apprenants qui ont juste dépassé (JD) le seuil sur la base d'une compréhension du GFP au [niveau 2 pour ODD 4.1.1(a), niveau 5 pour ODD 4.1.1(b) et niveau 8 pour ODD 4.1.1(c)].

Slide 97

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Étape 2 : Lisez attentivement le premier item de l'évaluation et envisagez les **connaissances et/ou compétences** nécessaires pour répondre correctement à l'item. Réfléchissez à ce qui rend l'item facile ou difficile (par exemple, la formulation de l'énoncé de l'item et la pertinence des options incorrectes, ou distrayeurs) et quels types d'erreurs peuvent être possibles ou raisonnables.

Comment écrit-on quatre-vingt-sept sous forme normalisée ?

A. 80
B. 87
C. 807
D. 870

Connaissances et compétences requises : Compter, lire et écrire des nombres entiers jusqu'à 100.
Énoncé de l'item : Il est clairement indiqué.
Distrayeurs de l'item : Les options A et C sont fortes.
Erreurs possibles : Les apprenants peuvent confondre sept et soixante-dix ou mal comprendre la valeur de position.

Slide 98

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Étape 3 : En vous basant sur la Tâche 2, sélectionnez le domaine, le concept, le sous-concept, l'énoncé des connaissances et/ou des compétences, et les GPL/GPD dans le GPF qui sont les plus pertinents pour l'item.

Domaine : Nombres et opérations
Concept : Nombres Entiers
Sous-concept : Identifier, compter et déterminer l'importance relative des nombres entiers
Connaissances ou compétences (norme de contenu) : Compter, lire et écrire des nombres entiers
GPL et GPD (normes de performance) :
Niveau scolaire : Niveau 2
Partiellement atteint : Lire et écrire les nombres entiers jusqu'à 30 en lettres et en chiffres.
Atteint : Lire et écrire les nombres entiers jusqu'à 100 en lettres et en chiffres.
Dépasse : N/D

Slide 99

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Étape 4 : En vous basant sur la compréhension des Étapes 1 à 3, suivez cette procédure :

- **Demandez si les apprenants JPA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item, c'est-à-dire si vous êtes raisonnablement sûr (≥ 67 % de chances, ou 2 des 3 apprenants JPA) ?**
 - Si « oui », entourez APJ et passez à l'item suivant.
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - Si « oui », entourez A et passez à l'item suivant.
 - Si « non », demandez si les apprenants D ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - » Si « oui », entourez JD et passez à l'item suivant.
 - » Si « non », entourez DA et passez à l'item suivant.

Slide 100

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Étape 2 : Lisez attentivement le premier item de l'évaluation et prenez en compte les **connaissances et/ou compétences** requises pour répondre correctement à l'item - n'oubliez pas de tenir compte de la pertinence du texte sur lequel l'item est basé pour le niveau scolaire. Réfléchissez à ce qui rend l'item facile ou difficile (par exemple, la formulation de l'énoncé de l'item et la pertinence des options incorrectes, ou distrayeurs) et quels types d'erreurs peuvent être possibles ou raisonnables.

Jabu avait un chien de compagnie. Il a emmené le chien dehors pour jouer. Le chien s'est endormi et s'est perdu. Jabu était triste. Après un moment, le chien est revenu. Jabu a ramené le chien à l'intérieur. Il a nourri le chien. Ensuite, le chien s'est endormi. Quand le chien s'est réveillé, Jabu l'a emmené dehors pour jouer à nouveau.

Qui possède un chien de compagnie ?

Connaissances et compétences requises : Récupérez une seule information explicite d'un texte contenu au niveau scolaire en effectuant une correspondance directe ou proche avec les mots.
Énoncé de l'item : Il est clairement indiqué.
Distrayeurs de l'item : Aucun autre nom dans le texte.
Erreurs possibles : Les apprenants pourraient ne pas savoir que « Jabu » est un prénom.

Slide 101

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Étape 3 : En vous basant sur la Tâche 2, sélectionnez le domaine, le concept, le sous-concept, l'énoncé des connaissances et/ou des compétences, et les GPL/GPD dans le GPF qui sont les plus pertinents pour l'item.

Domaine : Compréhension de la lecture
Concept : Récupérer des informations
Sous-concept : Récupérer des informations explicites dans un texte de niveau scolaire en utilisant une correspondance directe ou proche de mots
Connaissances ou compétences : Récupérer une seule information explicite à partir d'un texte contenu de niveau scolaire en utilisant une correspondance directe ou proche de mots
GPL et GPD (normes de performance) :
Niveau scolaire : Niveau 2
Partiellement atteint : Récupérez une seule information importante et explicite à partir d'un texte de niveau 2 en effectuant une correspondance directe ou proche des mots lorsque l'information requise est adjacente au mot correspondant et qu'il n'y a pas d'informations concurrentes. Cela sera généralement en réponse à une question de type « qui », « quoi », « quand » ou « où ».
Atteint : Récupérer une seule information explicite d'un texte de niveau 2 en ...
Dépasse : Récupérer une seule information explicite d'un texte de niveau 2 en ... lorsque les informations concurrentes sont limitées ...

Slide 102



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Étape 4 : En vous basant sur la compréhension des Étapes 1 à 3, suivez cette procédure :

- **Demandez si les apprenants JPA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item, c'est-à-dire si vous êtes raisonnablement sûr (≥ 67 % de chances, ou 2 des 3 apprenants JPA) ?**
 - Si « oui », entourez APJ et passez à l'item suivant.
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - Si « oui », entourez A et passez à l'item suivant.
 - Si « non », demandez si les apprenants D ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - » Si « oui », entourez JD et passez à l'item suivant.
 - » Si « non », entourez DA et passez à l'item suivant.

Slide 103



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Étape 2 : Estimez le nombre d'items que les apprenants JPA, JA et JD seraient capables de compléter dans la limite de temps (par exemple, les mots dans le passage de lecture orale que les apprenants essaieraient de lire en une minute).

Nbre de mots	Passage de lecture (Mot)	Série 1 : Nombre de mots que les apprenants essaieraient de lire en une minute			Évaluations individuelles et indépendantes de la série 1			Série 2 : Nombre de mots que les apprenants essaieraient de lire en une minute			Évaluations individuelles et indépendantes de la série 2				
		JPA	JA	JD	JPA	JA	JD	DA	JPA	JA	JD	JPA	JA	JD	DA
1	Kande	1	1	1	JPA	JA	JD	DA	1	1	1	JPA	JA	JD	DA
2	da	2	2	2	JPA	JA	JD	DA	2	2	2	JPA	JA	JD	DA
3	labbaharta	3	3	3	JPA	JA	JD	DA	3	3	3	JPA	JA	JD	DA
4	Detu	4	4	4	JPA	JA	JD	DA	4	4	4	JPA	JA	JD	DA
5	jukan	5	5	5	JPA	JA	JD	DA	5	5	5	JPA	JA	JD	DA

Slide 104



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Étape 3 : Lisez attentivement le premier mot ou la première question sur le [TA] et considérez les **connaissances et/ou compétences** requises pour lire correctement le mot ou répondre correctement à la question. Pensez à ce qui rend le mot ou la question facile ou difficile (par exemple, le type de connaissances et de compétences nécessaires, la formulation de la question) et quels types d'erreurs peuvent être possibles ou acceptables.

Item : Passage de décodage — Mot « Jabu »

Ce qui le rend facile ou difficile : C'est facile car il s'agit d'un mot simple et court suivant les règles orthographiques standard ; cela pourrait être difficile s'il ne s'agissait pas d'un nom courant.

Domaine : Décodage

Concept : Fluidité

Sous-concept : Lire ou signer un texte continu de niveau scolaire à un rythme et avec précision

Connaissances ou compétences : Dire ou signer avec fluidité un texte continu de niveau scolaire

Slide 105



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Étape 4 : En vous basant sur la Tâche 2, sélectionnez le domaine, le concept, le sous-concept, l'énoncé des connaissances et/ou des compétences, et les GPL/GPD dans le GPF qui sont les plus pertinents pour l'item.

Domaine : Décodage

Concept : Fluidité

Sous-concept : Lire ou signer un texte continu de niveau scolaire à un rythme et avec précision

Connaissances ou compétences : Dire ou signer avec fluidité un texte continu de niveau scolaire

GPL et GPD (normes de performance) :

Niveau scolaire : Niveau 2

Partiellement atteint : Lire ou signer avec précision un texte continu de niveau de 2, à un rythme lent selon les normes du pays en matière de fluidité pour la langue dans laquelle l'évaluation est administrée (par exemple, mot par mot).

Atteint : Lire ou signer avec précision un texte continu de niveau de 2, à un rythme conforme aux normes minimales du pays en matière de fluidité pour la langue dans laquelle l'évaluation est administrée.

Dépasse : Lire ou signer avec précision un texte continu de niveau de 2, à un rythme supérieur aux normes minimales du pays en matière de fluidité pour la langue dans laquelle l'évaluation est administrée.

Slide 106



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Étape 5 : En vous basant sur la compréhension des Étapes 1 à 3, suivez cette procédure :

- **Demandez si les apprenants JPA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item, c'est-à-dire si vous êtes raisonnablement sûr (≥ 67 % de chances, ou 2 des 3 apprenants JPA) ?**
 - Si « oui », entourez APJ et passez à l'item suivant.
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - Si « oui », entourez A et passez à l'item suivant.
 - Si « non », demandez si les apprenants D ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - » Si « oui », entourez JD et passez à l'item suivant.
 - » Si « non », entourez DA et passez à l'item suivant.

Slide 107

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Il y a des étapes légèrement différentes à suivre lors de l'évaluation des items comportant plusieurs points de note :

- **Étape 1 :** Identifiez et/ou conceptualisez trois apprenants juste partiellement atteint (JPA), trois apprenants juste atteint (JA) et trois apprenants juste dépasse (JD) en se basant sur une compréhension du GPF.
- **Étape 2 :** Lire attentivement l'item et prendre en compte les connaissances et/ou compétences requises pour répondre correctement à l'item [n'oubliez pas de considérer si le texte sur lequel l'item est basé est adapté au niveau de la classe.] Réfléchissez attentivement à ce que l'apprenant doit démontrer pour atteindre chaque point de note.
- **Étape 3 :** En vous basant sur la Tâche 2, sélectionnez le domaine, le concept, le sous-concept, l'énoncé des connaissances et/ou des compétences, et les GPU/GPD dans le GPF qui sont les plus pertinents pour l'item.

Slide 108

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

- **Étape 4 :** Demandez-vous si les apprenants JPA ayant un niveau de compétence minimal obtiendraient la première note?
 - Si « oui », entourez JPA pour la meilleure note et passez à la deuxième note
 - Si « non », demandez-vous si les apprenants JA ayant un niveau de compétence minimal obtiendraient la meilleure note?
 - Si « oui », entourez JA pour la meilleure note et passez à la deuxième note
 - Si « oui », demandez-vous si les apprenants JD ayant un niveau de compétence minimal obtiendraient la meilleure note?
 - » Si « oui », entourez JD et passez à la deuxième note
 - » Si « oui », entourez DA pour toutes les notes
- **Étape 5 :** Répétez l'étape 4 pour chaque note. À la fin, vous devriez avoir identifié combien de points de note vous pensez qu'un JPA, un JA et un JD pourraient obtenir sur l'item.

Slide 109

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Il y a quelques étapes légèrement différentes lors de l'examen d'items qui sont classés comme « sans correspondance » :

- **Étape 1 :** Identifiez et/ou conceptualisez trois apprenants juste partiellement atteint (JPA), trois apprenants juste atteint (JA) et trois apprenants juste dépasse (JD) en se basant sur une compréhension du GPF.
- **Étape 2 :** Lire attentivement l'item et prendre en compte les connaissances et/ou compétences requises pour répondre correctement à l'item, ainsi que la façon dont les JPA, JA et JD pourraient se comporter sur l'item

Slide 110

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Il y a quelques étapes légèrement différentes lors de l'examen d'items qui sont classés comme « sans correspondance » :

- **Étape 3 :** Demandez si les apprenants JPA ayant un niveau de compétence minimal répondraient correctement à l'item?
 - Si « oui », entourez JPA et passez à l'item suivant
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal répondraient correctement à l'item?
 - Si « oui », entourez JA et passer à l'item suivant
 - Si « non », demandez si les apprenants JD ayant un niveau de compétence minimal répondraient correctement à l'item?
 - » Si « oui », entourez JD et passez à l'item suivant
 - » Si « non », entourez DA et passez à l'item suivant

Slide 111

SÉRIE 1 : PROCÉDURE D'ÉVALUATION 

Organigramme pour l'évaluation des items avec des mots, des questions ou des items :

POUR CHAQUE MOT, QUESTION OU ITEMS :

```

graph TD
    A[Évaluez que 2 des 3 apprenants JPA seraient capables de lire le mot ou de répondre correctement à la question ou à l'item?] -- Non --> B[Évaluez que 2 des 3 apprenants JA seraient capables de lire le mot ou de répondre correctement à la question ou à l'item?]
    A -- Oui --> C[Entourez JPA.]
    B -- Non --> D[Évaluez que 2 des 3 apprenants JD seraient capables de lire le mot ou de répondre correctement à la question ou à l'item?]
    B -- Oui --> E[Entourez JA.]
    D -- Non --> F[Entourez DA et passez au mot, à la question ou à l'item suivant.]
    D -- Oui --> G[Entourez JD.]
  
```

REMARQUE : LORSQU'UN MOT, UNE QUESTION OU UN ITEM EST ENTOURÉ, PASSEZ AU MOT, À LA QUESTION OU À L'ITEM SUIVANT.

Slide 112

SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Organigramme pour l'évaluation des items avec des mots, des questions ou des items :

POUR CHAQUE MOT, QUESTION OU ITEMS :

Est-ce que 2 des 3 apprenants JA sont capables de lire le mot ou de répondre correctement à la question ?

OUI → Entourez OUI.

NON → Entourez Non, et passez au mot, à la question ou à l'item suivant.

NOTE: WHEN A CIRCLE IS MADE FOR A WORD, QUESTION, OR ITEM, PROCEED TO THE NEXT WORD, QUESTION, OR ITEM.

Slide 113

SÉRIE 1 : FORMULAIRE D'ÉVALUATION DES ITEMS

Instructions : Pour chaque item, entourez juste partiellement atteint (JPA), juste atteint (JA) ou juste dépasse (JD) le niveau de compétence minimale mondiale, selon si les apprenants ayant une compétence minimale à chaque niveau répondraient correctement à l'item (« oui »). Entourez « au-dessus du niveau de compétence minimale mondiale » (DA) pour les questions auxquelles même un apprenant JD ne serait pas en mesure de répondre correctement.

N° DE L'ITEM	SÉRIE 1				SÉRIE 2			
	JPA	JA	JD	DA	JPA	JA	JD	DA
1	JPA	JA	JD	DA	JPA	JA	JD	DA
2	JPA	JA	JD	DA	JPA	JA	JD	DA
3	JPA	JA	JD	DA	JPA	JA	JD	DA
4	JPA	JA	JD	DA	JPA	JA	JD	DA
5	JPA	JA	JD	DA	JPA	JA	JD	DA
6	JPA	JA	JD	DA	JPA	JA	JD	DA
7	JPA	JA	JD	DA	JPA	JA	JD	DA
8	JPA	JA	JD	DA	JPA	JA	JD	DA
9	JPA	JA	JD	DA	JPA	JA	JD	DA
10	JPA	JA	JD	DA	JPA	JA	JD	DA

Slide 114

SÉRIE 1 : FORMULAIRE D'ÉVALUATION DES ITEMS

Instructions : Pour chaque item, entourez juste partiellement atteint (JPA), juste atteint (JA) ou juste dépasse (JD) le niveau de compétence minimale mondiale, selon si les apprenants ayant une compétence minimale à chaque niveau répondraient correctement à l'item (« oui »). Entourez « au-dessus du niveau de compétence minimale mondiale » (DA) pour les questions auxquelles même un apprenant JD ne serait pas en mesure de répondre correctement.

Nbre de mots	Passage de lecture (Mot)	Série 1 : Nombre de mots que les apprenants réussissent de lire en une minute				Série 2 : Nombre de mots que les apprenants réussissent de lire en une minute									
		JPA	JA	JD	DA	JPA	JA	JD	DA						
1	écrite	1	1	1	JPA	JA	JD	DA	1	1	1	JPA	JA	JD	DA
2	les	2	2	2	JPA	JA	JD	DA	2	2	2	JPA	JA	JD	DA
3	aboyante	3	3	3	JPA	JA	JD	DA	3	3	3	JPA	JA	JD	DA
4	Deli	4	4	4	JPA	JA	JD	DA	4	4	4	JPA	JA	JD	DA
5	Iskan	5	5	5	JPA	JA	JD	DA	5	5	5	JPA	JA	JD	DA

Slide 115

SÉRIE 1 : CONSEILS UTILES POUR LA RÉALISATION DE L'ÉVALUATION DES ITEMS

- Bases la première série d'évaluations d'items sur les **directives suivantes** :
 - Effectuez les évaluations en fonction des **jugements individuels et indépendants** des items et du GPF.
 - Concentrez-vous sur le **contenu de l'item** par rapport aux énoncés de connaissances et/ou de compétences du GPF.
 - Prenez en considération la **difficulté de l'item**, y compris les erreurs possibles et raisonnables des apprenants.
 - Considérez plutôt ce que les apprenants « **feraient** » que ce qu'ils « **devraient** » pour réaliser une évaluation réaliste.

Slide 116

SÉRIE 1 : CALCUL DES RÉFÉRENCES MONDIALES

- Calculez les totaux des références initiales **pour chaque membre de l'équipe** :
 - **Partiellement atteint** = Total des « oui » dans la colonne JPA du formulaire d'évaluation
 - **Atteint** = Total des « oui » dans les colonnes JA et JA du formulaire d'évaluation
 - **Dépasse** = Total des « oui » dans les colonnes JPA, JA et JD du formulaire d'évaluation
- Calculez les moyennes des références globales initiales **pour l'équipe** :
 - **Partiellement atteint** = Moyenne des références atteintes partiellement pour tous les membres de l'équipe
 - **Atteint** = Moyenne des références atteintes pour tous les membres de l'équipe
 - **Dépasse** = Moyenne des références dépassées pour tous les membres de l'équipe

Slide 117

SÉRIE 1 : CALCUL DES RÉFÉRENCES MONDIALES

N° DE L'ITEM	SÉRIE 1				SÉRIE 2			
	JP	JM	JE	AE	JP	JM	JE	AE
1	JP	JM	JE	AE	JP	JM	JE	AE
2	JP	JM	JE	AE	JP	JM	JE	AE
3	JP	JM	JE	AE	JP	JM	JE	AE
4	JP	JM	JE	AE	JP	JM	JE	AE
5	JP	JM	JE	AE	JP	JM	JE	AE
6	JP	JM	JE	AE	JP	JM	JE	AE
7	JP	JM	JE	AE	JP	JM	JE	AE
8	JP	JM	JE	AE	JP	JM	JE	AE
9	JP	JM	JE	AE	JP	JM	JE	AE
10	JP	JM	JE	AE	JP	JM	JE	AE

Partiellement atteint = Total des « oui » dans la colonne JPA du formulaire d'évaluation = 4
 Atteint = Total des « oui » dans les colonnes JPA et JA du formulaire d'évaluation = 4 + 3 = 7
 Dépassé = Total des « oui » dans les colonnes JPA, JA et JD du formulaire d'évaluation = 4 + 3 + 2 = 9

PRÉSENTATION 13 – LES MEMBRES DE L'ÉQUIPE RÉALISENT LA TÂCHE D'ÉVALUATION SUR DES EXEMPLES PRATIQUES

Slide 118

TÂCHE 3 ACTIVITÉ

PRATIQUE DE LA MÉTHODE D'ANGOFF

Slide 119

ÉVALUATION DE LA PRATIQUE ITEM 1

Comment écrit-on quatre-vingt-sept sous forme normalisée ?

A. 80
B. 87
C. 807
D. 870

Domaine : Nombres et opérations
Concept : Nombres Entiers
Sous-concept : Identifier et compléter les nombres entiers, et identifier leur magnitude relative
Connaissances ou compétences (norme de contenu) : Compter, lire et écrire des nombres entiers

Qu'est-ce qui le rend facile ou difficile : les autres choix de réponses sont des distractions majeures, en particulier C.

GPL et GPD (norme de performance) :
GPD le plus bas pour répondre correctement — Partiellement atteint : Lire et écrire les nombres entiers jusqu'à 100 en lettres et en chiffres.
2 sur 3 des apprenants JPA répondraient-ils correctement à l'item ? ..
 Si oui, alors entourez JPA
 Si non, alors demandez en ce qui concerne JA ...

Slide 120

ÉVALUER L'ITEM DE PRATIQUE 2

Quelle est la différence de temps affichée entre ces deux horloges ?



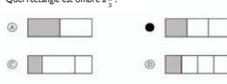
Domaine : Mesure
Concept : Heure
Sous-concept : Indiquer l'heure ET résoudre des problèmes impliquant l'heure
Connaissances ou compétences (norme de contenu) : Indiquer l'heure à l'aide d'une horloge analogique ET identifier ou résoudre des problèmes impliquant des équivalences entre différentes unités de temps

Qu'est-ce qui rend l'item facile/difficile ?
 Difficile—C'est un problème en deux étapes et les nombres ne sont pas indiqués sur les horloges.
Quel est le GPD le plus bas pour répondre correctement ?
 Niveau 3 Atteint—Indiquer l'heure à l'aide d'une horloge analogique à l'heure la plus proche ET résoudre des problèmes, y compris des problèmes du monde réel, impliquant un temps écoulé en heures. Ainsi, « dépasse » le niveau pour le niveau 2
2 sur 3 des apprenants JPA répondraient-ils correctement à l'item ? ..

Slide 121

ÉVALUER L'ITEM DE PRATIQUE 3

Quel rectangle est ombré $\frac{1}{3}$?



Domaine : Nombres et Opérations
Concept : Fractions
Sous-concept : Identifier et représenter des fractions à l'aide d'objets, d'images et de symboles, et identifier leur importance relative
Connaissances/compétences : Exprimer une représentation visuelle d'une fraction (image, objets) en notation fractionnaire

Qu'est-ce qui rend l'item facile ou difficile ?
 Difficile—comprendre que la portion ombrée doit être de 1/3 ombrée plutôt que simplement 1 des 3 parties doit être ombrée.
 C est une distraction majeure.
Quel est le GPD le plus bas pour répondre correctement ?
 Niveau 3 Partiellement atteint—Identifier des fractions unitaires quotidiennes représentées sous forme de dessins dans une notation fractionnaire. « Sans correspondance » pour le niveau 2
2 sur 3 des apprenants JPA répondraient-ils correctement à l'item ? ..

Slide 132



Slide 133

SÉRIE 1 : RÉSULTATS PAR ITEM

DISCUSSION SUR LA NOTATION DU NIVEAU 2
 Sur quels points sommes-nous en désaccord ?

Question : I2 + 7

Réponses : JPA : 1 JA : 6 JD : 5 DA : 0

« Atteint » : Additionner et soustraire dans la limite de 20
 « Dépasse » : Additionner et soustraire dans la limite de 30

Slide 134

SÉRIE 1 : LA COMPARAISON DES RÉSULTATS AVEC LA DIFFICULTÉ DES QUESTIONS

Difficulté de l'item :	Nombre de l'item		Valeur	
	Nombre	Valeur	Nombre	Valeur
1	0,17	21	0,40	
2	0,38	22	0,38	
3	0,50	23	0,36	
4	0,58	24	0,36	
5	0,79	25	0,67	
6	0,55	26	0,54	
7	0,59	27	0,69	
8	0,59	28	0,17	
9	0,59	29	0,56	
10	0,31	30	0,44	
11	0,47	31	0,71	
12	0,38	32	0,41	
13	0,47	33	0,68	
14	0,71	34	0,38	
15	0,40	35	0,39	
16	0,42	36	0,44	
17	0,34	37	0,39	
18	0,71	38	0,34	
19	0,43	39	0,83	
20	0,43	40	0,38	

Slide 135

SÉRIE 1 : RÉSULTATS UTILISANT LES DONNÉES D'IMPACT

Données d'impact :

Niveau de compétence minimum	Série 1 - Référence	Plage de notes	Pourcentage d'apprenants
En dessous du niveau partiellement atteint	N/D	0-12	44,5 %
Partiellement atteint	13	13-23	34,7 %
Répond	24	24-34	17,6 %
Dépasse	35	35-40	3,2 %
Total			100,0 %

PRÉSENTATION 16 – PRÉSENTATION DE LA DEUXIÈME SÉRIE 2 DE RÉFÉRENCES

Slide 136

PRÉSENTATION

TÂCHE 3 : ÉVALUATION COMPARATIVE ANGOFF SÉRIE 2

Slide 137

SÉRIE 2 DE LA PROCÉDURE DE NOTATION

- Effectuez la **deuxième série** d'évaluations en utilisant le même processus que pour la première série, c'est-à-dire la procédure en quatre étapes.
- Effectuez les notations des items de la série 2 sur la base des **orientations** suivantes :
 - Concentrez-vous sur le **contenu de l'item** par rapport aux énoncés de connaissances et/ou de compétences du GPF.
 - Maintenez la prise en compte de la **difficulté de l'item** comme base de jugement.
 - Apportez des **ajustements** à leurs notations sur la base de leurs jugements individuels et indépendants et du GPF.
 - Demandez-vous si vous êtes raisonnablement sûr que (2 apprenants sur 3) répondraient correctement à la question.
 - N'oubliez pas de prendre en compte la notion de « **serait** » plutôt que celle de « **devrait** » lorsque vous effectuez des évaluations réalistes.

Slide 138

PROCÉDURE ANGOFF : QUATRE ÉTAPES 

Étape 1 : Identifiez et/ou **conceptualisez** trois apprenants juste partiellement atteint (JPA), trois apprenants juste atteint (JA) et trois apprenants juste dépasse (JD) en se basant sur une compréhension du GPF.

Étape 2 : Lisez attentivement la première question de l'évaluation et réfléchissez aux **connaissances ou aux compétences** requises pour y répondre correctement [n'oubliez pas de vérifier si le texte sur lequel la question est basée est adapté au niveau]. Réfléchissez à ce qui rend l'item facile ou difficile (par exemple, la formulation de l'énoncé de l'item et la pertinence des options incorrectes, ou distracteurs) et quels types d'erreurs peuvent être possibles ou raisonnables.

Étape 3 : Sur la base de la tâche 2, sélectionnez le domaine, le concept, le sous-concept, le(s) énoncé(s) de connaissances et/ou de compétences et les GPL/GPD dans le GPF qui sont les plus pertinents pour l'item.

Slide 139

PROCÉDURE ANGOFF : QUATRE ÉTAPES

Étape 4 : En vous basant sur la compréhension des Étapes 1 à 3, suivez cette procédure :

- **Demandez si les apprenants JPA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item, c'est-à-dire si vous êtes raisonnablement sûr (≥ 67 % de chances, ou deux des trois apprenants JPA) ?**
 - Si « oui », entourez APJ et passez à l'item suivant.
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - Si « oui », entourez A et passez à l'item suivant.
 - Si « non », demandez si les apprenants D ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - » Si « oui », entourez JD et passez à l'item suivant.
 - » Si « non », entourez DA et passez à l'item suivant.

Slide 140

PROCÉDURE ANGOFF : CINQ ÉTAPES 

Étape 1 : Identifiez et/ou **conceptualisez** trois apprenants juste partiellement atteint (JPA), trois apprenants juste atteint (JA) et trois apprenants juste dépasse (JD) en se basant sur une compréhension du GPF.

Étape 2 : Estimez le nombre d'items que les apprenants JPA, JA et JD seraient capables de compléter dans la limite de temps (par exemple, les mots dans le passage de lecture orale que les apprenants essaieraient de lire en une minute).

Étape 3 : Lisez attentivement la première question de l'évaluation et réfléchissez aux **connaissances ou aux compétences** requises pour y répondre correctement. Réfléchissez à ce qui rend l'item facile ou difficile (par exemple, la formulation de l'énoncé de l'item et la pertinence des options incorrectes, ou distracteurs) et quels types d'erreurs peuvent être possibles ou raisonnables.

Étape 4 : Sur la base de la tâche 2, sélectionnez le domaine, le concept, le sous-concept, le(s) énoncé(s) de connaissances et/ou de compétences et les GPL/GPD dans le GPF qui sont les plus pertinents pour l'item.

Slide 141

PROCÉDURE ANGOFF : CINQ ÉTAPES 

Étape 5 : En vous basant sur la compréhension des Étapes 1 à 3, suivez cette procédure :

- **Demandez si les apprenants JPA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item, c'est-à-dire si vous êtes raisonnablement sûr (≥ 67 % de chances, ou deux des trois apprenants JPA) ?**
 - Si « oui », entourez APJ et passez à l'item suivant.
 - Si « non », demandez si les apprenants JA ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - Si « oui », entourez A et passez à l'item suivant.
 - Si « non », demandez si les apprenants D ayant un niveau de compétence minimal seraient capables de répondre correctement à l'item ?
 - » Si « oui », entourez JD et passez à l'item suivant.
 - » Si « non », entourez DA et passez à l'item suivant.

PRÉSENTATION 17– SÉRIE 2

Slide 142

TÂCHE 3 ACTIVITÉ

RÉALISER LA SÉRIE 2 DEL'ÉVALUATION D'ANGOFF

Slide 143



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Organigramme pour l'évaluation des items avec des mots, des questions ou des items :

POUR CHAQUE MOT, QUESTION OU ITEMS :

```

    graph TD
      Q1{Est-ce que 2 des 3 apprenants JPA seraient capables de lire le mot ou de répondre correctement à la question ou à l'item?}
      Q2{2 des 3 apprenants JA seraient-ils capables de lire le mot ou de répondre correctement à la question ou à l'item?}
      Q3{2 des 3 apprenants JD seraient-ils capables de lire le mot ou de répondre correctement à la question ou à l'item?}
      Q4{Entourez DA et passez au mot, à la question ou à l'item suivant.}

      Q1 -- Non --> A1[Entourez JPA.]
      Q1 -- Oui --> Q2
      Q2 -- Non --> Q3
      Q2 -- Oui --> A2[Entourez JA.]
      Q3 -- Non --> Q4
      Q3 -- Oui --> A3[Entourez JD.]
  
```

REMARQUE : LORSQU'UN MOT, UNE QUESTION OU UN ITEM EST ENTOURÉ, PASSEZ AU MOT, À LA QUESTION OU À L'ITEM SUIVANT.

Slide 144



SÉRIE 1 : PROCÉDURE D'ÉVALUATION

Organigramme pour l'évaluation des items avec des mots, des questions ou des items :

POUR CHAQUE MOT, QUESTION OU ITEMS :

```

    graph TD
      Q1{Est-ce que 2 des 3 apprenants JA seraient capables de lire le mot ou de répondre correctement à la question ou à l'item?}
      Q2{Entourez Non, et passez au mot, à la question ou à l'item suivant.}

      Q1 -- Non --> Q2
      Q1 -- Oui --> A1[Entourez Oui.]
  
```

REMARQUE : LORSQU'UN MOT, UNE QUESTION OU UN ITEM EST ENTOURÉ, PASSEZ AU MOT, À LA QUESTION OU À L'ITEM SUIVANT.

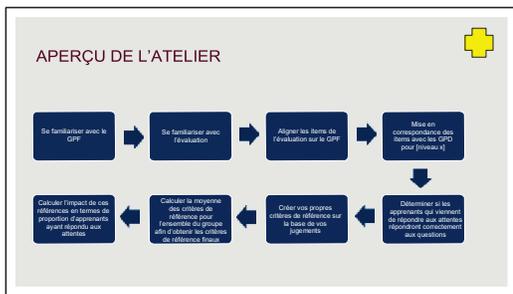
PRÉSENTATION 18– PRÉSENTATION DES RÉSULTATS DE LA SÉRIE 2

Slide 145

PRÉSENTATION

RÉVISION DES RÉSULTATS DE LA SÉRIE 2 DE L'ÉVALUATION D'ANGOFF

Slide 146



Slide 147

RÉSULTATS FINAUX ET PASSAGE D'UNE SÉRIE À L'AUTRE

Données d'impact :

Niveau de compétence minimum	SÉRIE 1			SÉRIE 2		
	Référence	Plage de notes	Pourcentage d'apprenants	Référence	Plage de notes	Pourcentage d'apprenants
En dessous du niveau partiellement atteint	N/A	0-12	44,5%	ND	0-14	50,9%
Partiellement atteint	13	13-23	34,7%	15	15-22	25,2%
Repère	24	24-34	17,6%	23	23-31	14,6%
Dépasse	35	35-40	3,2%	32	32-40	9,8%
Total			100,0%			100,0%

PRÉSENTATION 19– ÉVALUATION DE L'ATELIER

Slide 148



Slide 149

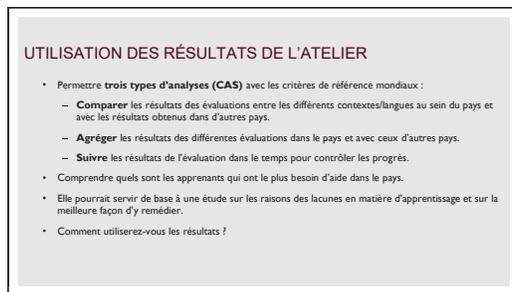
- INSTRUCTIONS POUR L'ÉVALUATION DE L'ATELIER**
- Vous allez maintenant remplir un **formulaire d'évaluation** pour nous faire part de vos opinions sur les aspects suivants de l'atelier :
 - Formation GPF
 - Formation à l'évaluation
 - Tâche d'alignement
 - Tâche de correspondance
 - Formation à l'établissement de liens entre les politiques
 - Résultats de la série 2

PRÉSENTATION 20 – CLÔTURE DE L'ATELIER ET LOGISTIQUE

Slide 150



Slide 151



Slide 152



Slide 153



Slide 154



ANNEXE H – FORMULAIRE D'ÉVALUATION DE L'ALIGNEMENT POUR LA TÂCHE 1

Pour mettre à jour ce formulaire, les animateurs doivent vérifier le nombre total de questions/d'items répertoriés à gauche et le modifier pour l'adapter aux besoins de l'évaluation utilisée. S'ils utilisent ce formulaire par voie électronique, les animateurs souhaiteront peut-être créer des menus déroulants conditionnels ou remplir automatiquement certaines colonnes.

Tableau 15 : modèle de formulaire d'évaluation de l'alignement

Question	Domaine	Référence du concept	Référence du sous-concept	Connaissance ou compétence	Correspondance	Domaine	Référence du concept	Référence du sous-concept	Connaissance ou compétence	Correspondance
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										

Ces colonnes ne sont requises qu'en cas de correspondance partielle. Vous pouvez les utiliser pour enregistrer tout autre domaine, concept et sous-concept liés à l'item.

ANNEXE I – FORMULAIRE D'ÉVALUATION D'ITEMS

Plusieurs exemples de formulaires d'évaluation d'items sont inclus ci-dessous. L'exemple de formulaire 1, comprenant le **tableau 14**, est un formulaire qui peut être utilisé pour définir trois critères de référence sur une **évaluation non chronométrée** de 45 items. Des items supplémentaires peuvent être ajoutés, selon les besoins. Pour adapter ce formulaire afin de définir un seul critère de référence, les animateurs doivent simplement supprimer les colonnes JPA et JD et renommer la colonne DA ADA (au-dessus d'atteint). D'autres exemples de formulaires sont inclus ci-dessous.

EXEMPLE DE FORMULAIRE 1. ÉVALUATION COMPRENANT 20 ITEMS OBJECTIFS (CHOIX MULTIPLES) :

3 apprenants JPA : _____
3 apprenants JA : _____
3 apprenants JD : _____

Nom du panéliste : _____
Code du panéliste : _____

Instructions : pour chaque item, entourez soit une compétence minimale juste partiellement atteinte (JPA), une compétence minimale juste atteinte (JA), une compétence minimale à peine dépassée (JD) ou une compétence minimale bien dépassée (AD).

Tableau 16 : Exemple de formulaire d'évaluation d'items pour les évaluations non chronométrées

Item n°	Prédictions individuelles et indépendantes de la série 1				Prédictions individuelles et indépendantes de la série 2			
	JPA	JA	JD	DA	JPA	JA	JD	DA
1	JPA	JA	JD	DA	JPA	JA	JD	DA
2	JPA	JA	JD	DA	JPA	JA	JD	DA
3	JPA	JA	JD	DA	JPA	JA	JD	DA
4	JPA	JA	JD	DA	JPA	JA	JD	DA
5	JPA	JA	JD	DA	JPA	JA	JD	DA
6	JPA	JA	JD	DA	JPA	JA	JD	DA
7	JPA	JA	JD	DA	JPA	JA	JD	DA
8	JPA	JA	JD	DA	JPA	JA	JD	DA
9	JPA	JA	JD	DA	JPA	JA	JD	DA
10	JPA	JA	JD	DA	JPA	JA	JD	DA
11	JPA	JA	JD	DA	JPA	JA	JD	DA
12	JPA	JA	JD	DA	JPA	JA	JD	DA
13	JPA	JA	JD	DA	JPA	JA	JD	DA
14	JPA	JA	JD	DA	JPA	JA	JD	DA
15	JPA	JA	JD	DA	JPA	JA	JD	DA
16	JPA	JA	JD	DA	JPA	JA	JD	DA
17	JPA	JA	JD	DA	JPA	JA	JD	DA
18	JPA	JA	JD	DA	JPA	JA	JD	DA
19	JPA	JA	JD	DA	JPA	JA	JD	DA
20	JPA	JA	JD	DA	JPA	JA	JD	DA
21	JPA	JA	JD	DA	JPA	JA	JD	DA
22	JPA	JA	JD	DA	JPA	JA	JD	DA
23	JPA	JA	JD	DA	JPA	JA	JD	DA

Item n°	Prédictions individuelles et indépendantes de la série 1				Prédictions individuelles et indépendantes de la série 2			
	JPA	JA	JD	DA	JPA	JA	JD	DA
24	JPA	JA	JD	DA	JPA	JA	JD	DA
25	JPA	JA	JD	DA	JPA	JA	JD	DA
26	JPA	JA	JD	DA	JPA	JA	JD	DA
27	JPA	JA	JD	DA	JPA	JA	JD	DA
28	JPA	JA	JD	DA	JPA	JA	JD	DA
29	JPA	JA	JD	DA	JPA	JA	JD	DA
30	JPA	JA	JD	DA	JPA	JA	JD	DA
31	JPA	JA	JD	DA	JPA	JA	JD	DA
32	JPA	JA	JD	DA	JPA	JA	JD	DA
33	JPA	JA	JD	DA	JPA	JA	JD	DA
34	JPA	JA	JD	DA	JPA	JA	JD	DA
35	JPA	JA	JD	DA	JPA	JA	JD	DA
36	JPA	JA	JD	DA	JPA	JA	JD	DA
37	JPA	JA	JD	DA	JPA	JA	JD	DA
38	JPA	JA	JD	DA	JPA	JA	JD	DA
39	JPA	JA	JD	DA	JPA	JA	JD	DA
40	JPA	JA	JD	DA	JPA	JA	JD	DA
41	JPA	JA	JD	DA	JPA	JA	JD	DA
42	JPA	JA	JD	DA	JPA	JA	JD	DA
43	JPA	JA	JD	DA	JPA	JA	JD	DA
44	JPA	JA	JD	DA	JPA	JA	JD	DA
45	JPA	JA	JD	DA	JPA	JA	JD	DA

L'exemple de formulaire 2 doit être utilisé avec **des questions à réponse construite/ouvertes** sur **des évaluations non chronométrées**. Les animateurs devront procéder à des ajustements en fonction du nombre de points possibles pour chaque question et du nombre total de questions (l'exemple ci-dessous ne comprend qu'un espace pour noter cinq questions). Il devrait y avoir une ligne incluse pour chaque valeur de point possible par question et chaque question. Des ajustements sont également nécessaires si les ateliers ne comportent que l'établissement d'un seul critère de référence (comme décrit ci-dessus).

EXEMPLE DE FORMULAIRE 2. ÉVALUATION COMPRENANT CINQ ITEMS OUVERTS

(L'item 1 a un score de 2 points, les items 2 et 3 ont un score de 4 points, l'item 4 a un score de 3 points et l'item 5 a un score de 5 points).

3 apprenants JPA : _____
3 apprenants JA : _____
3 apprenants JD : _____

Nom du panéliste : _____
Code du panéliste : _____

Instructions : pour chaque item, entourez soit une compétence minimale juste partiellement atteinte (JPA), une compétence minimale juste atteinte (JA), une compétence minimale à peine dépassée (JD) ou une compétence minimale bien dépassée (AD).

Tableau 17 : exemple de formulaire d'évaluation d'items pour les évaluations avec des questions à réponse construite

Item n°	Point	Prédictions individuelles et indépendantes de la série 1				Prédictions individuelles et indépendantes de la série 2			
		JPA	JA	JD	DA	JPA	JA	JD	DA
1	1-1	JPA	JA	JD	DA	JPA	JA	JD	DA
1	1-2	JPA	JA	JD	DA	JPA	JA	JD	DA
2	2-1	JPA	JA	JD	DA	JPA	JA	JD	DA
2	2-2	JPA	JA	JD	DA	JPA	JA	JD	DA
2	2-3	JPA	JA	JD	DA	JPA	JA	JD	DA
2	2-4	JPA	JA	JD	DA	JPA	JA	JD	DA
3	3-1	JPA	JA	JD	DA	JPA	JA	JD	DA
3	3-2	JPA	JA	JD	DA	JPA	JA	JD	DA
3	3-3	JPA	JA	JD	DA	JPA	JA	JD	DA
3	3-4	JPA	JA	JD	DA	JPA	JA	JD	DA
4	4-1	JPA	JA	JD	DA	JPA	JA	JD	DA
4	4-2	JPA	JA	JD	DA	JPA	JA	JD	DA
4	4-3	JPA	JA	JD	DA	JPA	JA	JD	DA
5	5-1	JPA	JA	JD	DA	JPA	JA	JD	DA
5	5-2	JPA	JA	JD	DA	JPA	JA	JD	DA
5	5-3	JPA	JA	JD	DA	JPA	JA	JD	DA
5	5-4	JPA	JA	JD	DA	JPA	JA	JD	DA
5	5-5	JPA	JA	JD	DA	JPA	JA	JD	DA

L'exemple de formulaire 3 fournit un exemple de formulaire pouvant être utilisé pour des **évaluations chronométrées**. L'exemple provient d'un atelier consacré à l'établissement de liens entre les politiques axé sur l'établissement de critères de référence pour l'EGRA. Il existe des colonnes supplémentaires nécessaires pour les évaluations chronométrées, car les panélistes doivent d'abord déterminer combien d'items/de mots un apprenant tentera dans le temps imparti, puis déterminer si les apprenants répondront à chacun des items/liront chacun de ces mots correctement ou non (uniquement jusqu'au nombre que le panéliste détermine que les apprenants de ce niveau de performance tenteront). Par exemple, si un panéliste dit qu'un apprenant JPA tentera 10 mots, lors de la deuxième étape du processus d'évaluation des évaluations chronométrées, il évaluera uniquement si l'apprenant répondrait correctement à ces dix premiers mots (p. ex., jusqu'au mot Wata, dans l'exemple ci-dessous). À l'instar des formulaires ci-dessus, ce formulaire doit être ajusté en fonction du nombre total d'items ainsi que du nombre de critères qui seront définis lors de l'atelier. Une autre différence avec ce formulaire est qu'au lieu d'inclure simplement le numéro de l'item, dans ce cas, il inclut l'item réel (dans ce cas, « mot » dans un passage de lecture). Les items pourraient également être ajoutés aux formulaires ci-dessus pour plus de clarté. Cela n'est généralement nécessaire que lorsque les numéros d'items ne sont pas clairement indiqués sur l'évaluation.

EXEMPLE DE FORMULAIRE 3. SOUS-TÂCHE DE FLUIDITÉ EN LECTURE ORALE AVEC 35 MOTS ET 5 ITEMS DE COMPRÉHENSION À LA LECTURE

3 apprenants JPA : _____
3 apprenants JA : _____
3 apprenants JD : _____

Nom du panéliste : _____
Code du panéliste : _____

Instructions : pour chaque item, entourez soit une compétence minimale juste partiellement atteinte (JPA), une compétence minimale juste atteinte (JA), une compétence minimale à peine dépassée (JD) ou une compétence minimale bien dépassée (AD).

Tableau 18 : Exemple de formulaire d'évaluation d'items pour une évaluation de lecture chronométrée (en haoussa)

Nbre de mots	Passage de lecture (mot)	Série 1 : nombre de mots que les apprenants tenteraient de lire en une minute			Évaluations individuelles et indépendantes de la série 1				Série 2 : nombre de mots que les apprenants tenteraient de lire en une minute			Évaluations individuelles et indépendantes de la série 2			
		JPA	JA	JD	JPA	JA	JD	DA	JPA	JA	JD	JPA	JA	JD	DA
1	Kande	1	1	1	JPA	JA	JD	DA	1	1	1	JPA	JA	JD	DA
2	da	2	2	2	JPA	JA	JD	DA	2	2	2	JPA	JA	JD	DA
3	abokiyarta	3	3	3	JPA	JA	JD	DA	3	3	3	JPA	JA	JD	DA
4	Delu	4	4	4	JPA	JA	JD	DA	4	4	4	JPA	JA	JD	DA
5	sukan	5	5	5	JPA	JA	JD	DA	5	5	5	JPA	JA	JD	DA
6	tafi	6	6	6	JPA	JA	JD	DA	6	6	6	JPA	JA	JD	DA
7	Makaranta	7	7	7	JPA	JA	JD	DA	7	7	7	JPA	JA	JD	DA
8	tare	8	8	8	JPA	JA	JD	DA	8	8	8	JPA	JA	JD	DA
9	kullum.	9	9	9	JPA	JA	JD	DA	9	9	9	JPA	JA	JD	DA
10	Wata	10	10	10	JPA	JA	JD	DA	10	10	10	JPA	JA	JD	DA
11	rana	11	11	11	JPA	JA	JD	DA	11	11	11	JPA	JA	JD	DA
12	Kande	12	12	12	JPA	JA	JD	DA	12	12	12	JPA	JA	JD	DA
13	ta	13	13	13	JPA	JA	JD	DA	13	13	13	JPA	JA	JD	DA
14	zo	14	14	14	JPA	JA	JD	DA	14	14	14	JPA	JA	JD	DA
15	da	15	15	15	JPA	JA	JD	DA	15	15	15	JPA	JA	JD	DA
16	aiki	16	16	16	JPA	JA	JD	DA	16	16	16	JPA	JA	JD	DA
17	daga	17	17	17	JPA	JA	JD	DA	17	17	17	JPA	JA	JD	DA
18	makaranta.	18	18	18	JPA	JA	JD	DA	18	18	18	JPA	JA	JD	DA
19	Delu	19	19	19	JPA	JA	JD	DA	19	19	19	JPA	JA	JD	DA
20	ta	20	20	20	JPA	JA	JD	DA	20	20	20	JPA	JA	JD	DA
21	taimaka	21	21	21	JPA	JA	JD	DA	21	21	21	JPA	JA	JD	DA
22	mata.	22	22	22	JPA	JA	JD	DA	22	22	22	JPA	JA	JD	DA
23	Kande	23	23	23	JPA	JA	JD	DA	23	23	23	JPA	JA	JD	DA
24	ta	24	24	24	JPA	JA	JD	DA	24	24	24	JPA	JA	JD	DA
25	samu	25	25	25	JPA	JA	JD	DA	25	25	25	JPA	JA	JD	DA
26	yabo	26	26	26	JPA	JA	JD	DA	26	26	26	JPA	JA	JD	DA
27	a	27	27	27	JPA	JA	JD	DA	27	27	27	JPA	JA	JD	DA
28	ajinsu.	28	28	28	JPA	JA	JD	DA	28	28	28	JPA	JA	JD	DA
29	Kande	29	29	29	JPA	JA	JD	DA	29	29	29	JPA	JA	JD	DA
30	da	30	30	30	JPA	JA	JD	DA	30	30	30	JPA	JA	JD	DA
31	Delu	31	31	31	JPA	JA	JD	DA	31	31	31	JPA	JA	JD	DA
32	Sun	32	32	32	JPA	JA	JD	DA	32	32	32	JPA	JA	JD	DA
33	ji	33	33	33	JPA	JA	JD	DA	33	33	33	JPA	JA	JD	DA
34	dadi	34	34	34	JPA	JA	JD	DA	34	34	34	JPA	JA	JD	DA
35	sosai.	35	35	35	JPA	JA	JD	DA	35	35	35	JPA	JA	JD	DA
Total															

La deuxième partie de l'exemple de formulaire 3 peut également être utilisée avec des évaluations chronométrées, telles que EGRA/EGMA, ou d'autres évaluations avec des questions conditionnelles. Cet exemple provient de la sous-tâche de compréhension écrite de l'EGRA. La sous-tâche de compréhension écrite de l'EGRA exige que les recenseurs lisent uniquement aux apprenants le nombre de questions de compréhension écrite qui correspondent au nombre de mots que l'apprenant a tenté, comme indiqué dans la colonne « condition » du formulaire ci-dessous. En tant que tel, il est important que lors de l'évaluation d'une sous-tâche, telle que la sous-tâche de compréhension écrite de l'EGRA, les panélistes prennent en compte le nombre qu'ils estiment que les apprenants d'un niveau de performance spécifique auraient tenté. Ainsi, en développant l'exemple ci-dessus, cela signifierait que si un panéliste estime que les apprenants JPA liraient 10 mots dans le passage, alors ces apprenants JPA ne se verraient poser que la première question du tableau ci-dessous (selon les critères énumérés dans la section « condition »). Ainsi, ils ne devraient évaluer la première question que par oui/non pour les apprenants JPA. Ce formulaire devra être adapté en fonction du nombre d'items, des conditions de ces items, des items eux-mêmes et du nombre de critères de référence.

Tableau 19 : Exemple de formulaire d'évaluation d'items pour les questions de compréhension de lecture conditionnelles (en haoussa)

Item n°	Condition	Questions	Évaluations individuelles et indépendantes de la série 1				Évaluations individuelles et indépendantes de la série 2			
			JPA	JA	JD	DA	JPA	JA	JD	DA
1	≤ 9 mots tentés	Su waye abokan juna? <i>{Kande da Delu}</i>	JPA	JA	JD	DA	JPA	JA	JD	DA
2	≤ 18 mots tentés	Ina suke tafiya kullum? <i>{Makaranta}</i>	JPA	JA	JD	DA	JPA	JA	JD	DA
3	≤ 22 mots tentés	Me Kande ta zo da shi daga makaranta? <i>{Aiki}</i>	JPA	JA	JD	DA	JPA	JA	JD	DA
4	≤ 28 mots tentés	Wa ya taimaka wa Kande? <i>{Delu}</i>	JPA	JA	JD	DA	JPA	JA	JD	DA
5	≤ 35 mots tentés	Me ya faru a ajin su Kande? <i>{Kande ta Samu yabo/ yabo}</i>	JPA	JA	JD	DA	JPA	JA	JD	DA
Total										

ANNEXE J – STATISTIQUES DE PRÉCISION, D'EXACTITUDE ET DE COHÉRENCE

PANÉLISTES ATYPIQUES

Pour identifier les panélistes dont les notes sont clairement aberrantes par rapport à l'autre membre du panel, le modèle Interquartile ou méthode de Tukey sera utilisé.

Toutes les notes de coupure inférieures à $Q1 - (Q3 - Q1)$ ou supérieures à $Q1 + K(Q3 - Q1)$ seront considérées comme des valeurs aberrantes.

Où,

Q_1 = quartile inférieur

Q_3 = quartile supérieur

$K = 1,5$

$K = 1,5$ a été proposé comme multiplicateur par Tukey (1977) et a été principalement utilisé depuis.

COHÉRENCE ENTRE LES ÉVALUATEURS

La cohérence entre les évaluateurs est calculée à l'aide de la formule généralisée de Ferdous & Plake (2005) pour plusieurs critères de référence. La procédure est basée sur la différence absolue entre les réponses de deux panélistes pour toutes les paires possibles de panélistes. Cet indice peut être calculé à la fois au niveau des items (c'est-à-dire pour les évaluations des items par les panélistes) et pour l'ensemble du test. La cohérence entre les évaluateurs pour un item i est définie comme la proportion du total des cohérences observées par rapport au nombre total de cohérences possibles. La cohérence totale observée est définie par la somme des différences absolues de toutes les paires possibles de réponses des panélistes.

La cohérence entre les évaluateurs pour l'item i est,

$$I_i = 1 - \frac{TOI_i}{TI} \quad (4)$$

$$TOI_i = \frac{z!}{\sum_{a,b=1}^z a \neq b} |R_{ai} - R_{bi}| \quad (5)$$

$$TI = d * \frac{z!}{2(z-2)!} \quad (6)$$

Où,

I_i = Cohérence entre les évaluateurs pour l'item i . Un nombre élevé (0,80 et plus) indique une cohérence élevée et un nombre faible indique une cohérence faible

TOI_i = Incohérence totale entre les évaluateurs observée pour l'item i

TI = Incohérence totale possible entre les évaluateurs pour chaque item

- Z = Nombre de panélistes dans l'étude normative
- R_{ai} = Réponse du panéliste a à l'item i ; $k = 1, 2, 3, 4$ (1 = atteint partiellement, 4 = dépasse nettement) ou 1, 2 (1 = atteint, 2 = dépasse le niveau pour un critère de référence)
- R_{bi} = Réponse du panéliste b à l'item i ; $k = 1, 2, 3, 4$ (1 = atteint partiellement, 4 = dépasse nettement) ou 1, 2 (1 = en dessous du niveau, 2 = au-dessus d'un critère de référence)
- d = Différence absolue maximale possible entre les notes de deux juges.

S'il y a quatre catégories de niveaux de réussite, un juge peut attribuer une note de 1 (atteint partiellement) à l'item et l'autre juge peut donner une note de 4 (dépasse nettement les compétences minimales) ; ainsi, la différence absolue maximale possible est de 3. S'il y a deux catégories de niveaux de réussite, un juge peut attribuer une note de 1 (atteint) à l'item et l'autre juge peut donner une note de 2 (dépasse le niveau) ; ainsi, la différence absolue maximale possible est de 1.

La cohérence globale pour un nombre n d'items du test parmi tous les panélistes est la suivante :

$$I = n^{-1} \sum_i^n I_i = 1 \quad I_i \quad (7)$$

Comment calculer la cohérence entre les évaluateurs

Calculer la cohérence entre les évaluateurs pour un item et l'ensemble de l'évaluation.

Étape 1 : calculer l'incohérence totale possible entre les évaluateurs.

- i. Calculer la factorielle du nombre de panélistes.
- ii. Calculer la factorielle de deux fois le nombre de panélistes moins deux.
- iii. Diviser les résultats de la sous-étape 1 par le résultat de la sous-étape 2.
- iv. Multiplier la différence absolue maximale possible entre les notes de deux juges par le résultat de la sous-étape 3. Ce résultat correspond à l'incohérence totale possible entre les évaluateurs.

Étape 2 : calculer la cohérence entre les évaluateurs pour un item.

- i. Prendre la valeur absolue de la différence de notes entre chaque panéliste.
- ii. Additionner toutes les valeurs absolues. Le résultat correspond à l'incohérence totale observée entre les évaluateurs pour l'item.
- iii. Diviser l'incohérence totale entre les évaluateurs observée pour l'item par l'incohérence totale possible entre évaluateurs. Le résultat est la cohérence entre les évaluateurs pour l'item.
- iv. Répéter les sous-étapes 1 à 3 pour chaque item de l'évaluation.

Étape 3 : calculer la cohérence entre les évaluateurs pour l'évaluation.

- i. Additionner l'incohérence entre les évaluateurs de chaque item.
- ii. Diviser la somme par le nombre d'items de l'évaluation. Le résultat correspond à la cohérence entre les évaluateurs.

ERREUR TYPE (SE)

L'erreur type (SE) est calculée pour chaque critère de référence séparément à l'aide des formules suivantes :

$$SE \text{ (Partiellement atteint les critères)} = \frac{SD_{(1)}}{\sqrt{z-1}} \quad (8)$$

$$SE \text{ (Atteint les critères)} = \frac{SD_{(2)}}{\sqrt{z-1}} \quad (9)$$

$$SE \text{ (dépasse le critère de compétence minimum)} = \frac{SD_{(3)}}{\sqrt{z-1}} \quad (10)$$

Où,

$SD_{(1)}$ = écart type du critère de référence partiellement atteint pour tous les panélistes z

$SD_{(2)}$ = écart type du critère de référence atteint pour tous les panélistes z

$SD_{(3)}$ écart type du critère de référence dépasse la compétence minimale pour tous les z panélistes

z = Nombre total de panélistes

Comment calculer l'erreur type des critères de référence

Calculer la SE pour un critère de référence.

- 1) Prendre les repères de tous les panélistes et calculer l'écart type des critères des panélistes.
- 2) Soustraire 1 au nombre total de panélistes.
- 3) Calculer la racine carrée du résultat de l'étape 2.
- 4) Diviser le résultat de l'étape 1 par les résultats de l'étape 3. Le résultat est la SE pour ce critère de référence.
- 5) Répéter les étapes 1 à 4 si nécessaire pour chaque critère de référence.

INTERVALLES DE CONFIANCE

L'intervalle de confiance de 95 % pour chaque critère de référence est calculé à l'aide de la formule suivante :

$$\text{Confidence Interval} = B_i \pm 1.96 \times SE_i$$

Où,

B_i = le critère de référence pertinent pour le MPL

SE_i = l'erreur type pertinente pour le MPL

ANNEXE K – MODÈLE DE LETTRE D'INVITATION POUR LES OBSERVATEURS

Cette annexe comprend un modèle de lettre destinée aux observateurs du gouvernement/de l'agence d'évaluation et d'autres organisations intéressées. Tous les détails à renseigner sont indiqués entre parenthèses. La lettre doit être modifiée selon les besoins pour s'adapter au contexte.

[Date]

[Nom]

[Rôle]

[Agence]

[Adresse/lieu]

Invitation à un atelier consacré à l'établissement de liens entre les politiques pour mesurer les résultats d'apprentissage mondiaux

Cher [Nom],

Dans la poursuite des Objectifs de développement durable sur l'éducation (ODD 4.1.1), [l'Évaluation nationale/régionale ou internationale] a décidé de procéder à l'utilisation d'une méthode de rapport mondiale appelée « Établissement de liens entre les politiques pour mesurer les résultats d'apprentissage mondiaux » (appelé Établissement de liens entre les politiques dans tout le document). Cette méthode permet aux pays/agences d'évaluation de déterminer si leurs apprenants atteignent les compétences minimales mondiales en lecture et en mathématiques, selon les ODD

4.1.1. [L'USAID utilise des indicateurs similaires pour ses rapports mondiaux].

Grâce à l'établissement de liens entre les politiques, les pays/agences d'évaluation relient leurs évaluations nationales à une échelle de rapport mondiale commune à l'aide de critères de référence. L'établissement de critères nécessite des jugements sur les performances des apprenants par des panels d'experts en matière de programme d'études et des enseignants. Les critères permettront de déterminer le pourcentage d'apprenants atteignant un niveau de compétence minimal en lecture et en mathématiques.

[Le Pays/l'Agence d'évaluation] prévoit d'organiser un atelier consacré à l'établissement de liens entre les politiques [en personne/à distance] du [date de début] au [date de fin]. **L'inscription aura lieu à [heure] le [date].** L'atelier se concentrera sur le lien entre [Nom(s) de l'évaluation] et l'ODD 4.1.1 pour les [Classes X et Y]. Il y aura [nombre X] panels, [un pour la langue d'évaluation des notes X et un pour la langue d'évaluation des notes Y – peut également en inclure plus de deux]. Les panélistes engloberont des maîtres enseignants et des experts en programmes d'études, et ils seront guidés à travers un processus systématique qui implique l'examen du matériel d'évaluation et l'établissement de critères de référence pour [la(les) langue(s) d'évaluation des notes].

Jusqu'à [nombre] administrateurs de [Agence] sont invités à participer en tant qu'observateurs. La participation à l'atelier offrira aux administrateurs sélectionnés l'occasion de : 1) s'appuyer sur les résultats de l'atelier sur le cadre national de lecture, 2) d'en apprendre davantage sur la méthode d'établissement de liens entre les politiques du monde entier pour l'établissement de rapports au regard de l'ODD 4.1.1, 3) et fournir des informations générales et leur expérience pour que l'établissement de liens entre les politiques puissent être étendus [dans/avec le pays/l'agence d'évaluation] aux évaluations pour d'autres niveaux scolaires, domaines et langues.

Nom de l'activité	Date d'arrivée	Date de départ	Lieu
[Nom de l'atelier]	[Date] Inscription à [Heure]	[Date] La dernière session se termine à [Heure]	[Lieu] pour l'atelier et [Hôtel] pour l'hébergement des participants de l'extérieur de la ville

[Détails logistiques, p. ex, qui couvrira les frais de transport, d'hébergement, les indemnités journalières, les déjeuners]

Si vous avez des questions ou avez besoin de précisions supplémentaires, veuillez contacter [Nom] par téléphone [numéro]. Merci de bien vouloir confirmer votre participation avant le [Date]. Votre participation à cet atelier est cruciale et nous espérons collaborer avec vous.

Cordialement,

[Nom et titre]

ANNEXE L – MODÈLE DE LETTRE D’INVITATION POUR LES PANÉLISTES DES ATELIERS

Cette annexe comprend un modèle de lettre pour les panélistes, à la fois experts en programmes d’études et enseignants. Tous les détails qui doivent être remplis sont inclus entre parenthèses. La lettre doit être modifiée au besoin pour s’adapter au contexte.

[Date]

Cher [Nom],

Invitation à un atelier consacré à l’établissement de liens entre les politiques

Dans la poursuite des Objectifs de développement durable sur l’éducation (ODD 4.1.1), [l’Évaluation nationale/régionale ou internationale] a décidé de procéder à l’utilisation d’une méthode de rapport mondiale appelée « Établissement de liens entre les politiques pour mesurer les résultats d’apprentissage mondiaux » (appelé Établissement de liens entre les politiques dans tout le document). Cette méthode permet aux pays/agences d’évaluation de déterminer si leurs apprenants atteignent les compétences minimales mondiales en lecture et en mathématiques, selon les ODD 4.1.1.

Grâce à l’établissement de liens entre les politiques, les pays/agences d’évaluation relieront leurs évaluations nationales à une échelle de rapport mondiale commune à l’aide de critères de référence. L’établissement de critères de référence nécessite l’avis de panels d’enseignants.

[Le Pays/l’Agence d’évaluation] prévoit d’organiser un atelier consacré à l’établissement de liens entre les politiques [en personne/à distance] du **[date de début] au [date de fin]. L’inscription aura lieu à [heure] le [date].** L’atelier se concentrera sur le lien entre [Nom(s) de l’évaluation] et l’ODD 4.1.1 pour les [Classes X et Y]. Il y aura [nombre X] panels, [un pour la langue d’évaluation des notes X et un pour la langue d’évaluation des notes Y – peut également en inclure plus de deux]. Les panélistes comprendront des maîtres enseignants et des experts en programmes d’études, et ils seront guidés à travers un processus systématique qui implique l’examen du matériel d’évaluation et l’établissement de critères de référence pour [la(les) langue(s) d’évaluation des notes].

[Le ministère gouvernemental/l’agence d’évaluation] a besoin d’un total de [nombre de panélistes] pour participer à l’atelier, y compris [nombre X du lieu, avec une expérience en matière de niveau scolaire, de matière et de langue d’évaluation ; Numéro Y de . . .]. À ce titre, [ministère gouvernemental/agence d’évaluation] souhaite vous inviter à participer à l’atelier.

La participation à l’atelier offrira une précieuse opportunité d’apprentissage aux panélistes sélectionnés, qui acquerront une meilleure compréhension des normes internationales en matière de performance des apprenants.

Nom de l’activité	Date d’arrivée	Date de départ	Lieu
[Nom de l’atelier]	[Date] Inscription à [Heure]	[Date] La dernière session se termine à [Heure]	[Lieu] pour l’atelier et [Hôtel] pour l’hébergement des participants de l’extérieur de la ville

[Détails logistiques, p. ex, qui couvrira les frais de transport, d’hébergement, les indemnités journalières, les déjeuners]

Si vous avez des questions ou avez besoin de précisions supplémentaires, veuillez contacter [Nom] par téléphone [numéro]. Merci de bien vouloir confirmer votre participation avant le [Date]. Si vous décidez de participer, nous vous demandons de réaliser l'activité préalable à l'atelier détaillée dans la pièce jointe à cette lettre avant l'atelier. Votre participation à cet atelier est cruciale et nous espérons que vous nous rejoindrez.

Cordialement,

[Nom et titre]

ANNEXE M – INFORMATIONS DÉMOGRAPHIQUES DES PANÉLISTES

Les animateurs doivent mettre à jour ce formulaire pour refléter les distinctions géographiques (en particulier, la région et le district) qui doivent être suivies afin d'assurer une représentativité appropriée du panel pour l'atelier et doivent ajouter tout autre détail nécessaire au rapport. Ils peuvent également souhaiter créer un formulaire électronique qui permette une saisie plus facile des données.

Groupe de sujets : 1) Lecture
2) Niveau scolaire

en mathématiques : _____

Langue : _____

Nom : _____

Profession : _____

Région où vous enseignez/travaillez : _____

Département où vous enseignez/travaillez : _____

Courriel : _____

Téléphone portable : _____

Sexe : 1) Femme
2) Homme

Origine ethnique (le cas échéant) : _____

Niveau d'études : _____

Années d'expérience/de savoir-faire : _____

Années d'enseignement/de travail avec un niveau et une matière pertinents : _____

Organisation/affiliation professionnelle (p. ex, école, ministère, etc.) : _____

Formation(s) antérieure(s) en lecture/mathématiques (réponse uniquement pour le sujet pour lequel vous êtes panéliste :

1) Non
2) Oui

Expérience dans l'enseignement auprès d'apprenants en situation de handicap :

1) Non
2) Oui

Expérience professionnelle avec les populations touchées par les conflits et les crises :

- 1) Non
- 2) Oui

Langue maternelle : _____

Langue(s) utilisée(s) pour l'enseignement en classe (pour les enseignants uniquement) : _____

ANNEXE N – STATISTIQUES AVANT L’ATELIER

L’analyste de données et/ou l’animateur principal doivent calculer les statistiques suivantes avant l’atelier consacré à l’établissement de liens entre les politiques. La méthode utilisée pour calculer ces statistiques varie et dépend du modèle d’évaluation utilisé. Pour les évaluations où les statistiques de la théorie classique des tests sont valides (par exemple, lorsque la cohorte entière ou un échantillon représentatif d’apprenants répond à tous les items de l’évaluation), l’approche de la théorie classique des tests (CTT) doit être suivie. Lorsqu’un plan d’échantillonnage complexe est utilisé avec une analyse de la théorie des réponses aux items (IRT), l’approche IRT doit alors être suivie.

DIFFICULTÉ DE L’ITEM

La difficulté de l’item informe les animateurs et les panélistes sur la difficulté d’un item en fonction de la façon dont les apprenants ont réalisé cet item lors de l’itération la plus récente de l’évaluation. L’analyste de données doit calculer le niveau de difficulté des items empiriques en suivant les étapes suivantes :

Théorie classique des tests

1. Calculer le niveau de difficulté empirique des items pour chacun d’eux en calculant la proportion d’apprenants qui réussissent un item. Ce sont les informations que vous présenterez aux panélistes entre la série 1 de notation de référence 1 et la série 2.

Théorie de la réponse aux items

1. Selon le modèle IRT utilisé, trois paramètres clés sont généralement rapportés pour un item : la difficulté de discrimination et la difficulté de deviner. Si un modèle IRT à 1 paramètre est utilisé, ou si l’on peut supposer que les items ont été alternés de manière aléatoire et avec un taux d’exposition approximativement égal dans la cohorte de tests, la difficulté des items peut être calculée en utilisant la même approche que le CTT.
2. Lorsqu’un modèle à 2 ou 3 paramètres est utilisé et que l’on craint qu’il existe des différences dans les capacités des apprenants qui ont suivi chacun des items, le paramètre de difficulté de l’item IRT doit être calculé avec une probabilité de réponse de 0,50. Lors du partage de ces informations avec le panéliste, il convient de veiller à expliquer l’échelle IRT, car la difficulté des items calculée à partir de l’IRT est moins intuitive que pour le CTT.

DIFFUSIONS DES DONNÉES

L’analyste de données peut préparer des informations sur la diffusion des données de l’itération la plus récente de l’évaluation liée au GPF et à l’ODD 4.1.1 avant l’atelier, bien que les données ne soient pas nécessaires avant le jour 4, entre les évaluations des séries 1 et 2. Se préparer à l’avance permet de gagner une étape dans le calendrier généralement contraint de l’atelier.

Théorie classique des tests

Pour préparer les distributions, l’analyste de données analysera le nombre et le pourcentage d’apprenants ayant passé l’évaluation ayant reçu une note globale de zéro, de même pour les apprenants ayant reçu une note globale de un, et ainsi de suite jusqu’à la note la plus élevée possible à l’évaluation. Ils auront recours à ces informations pour préparer un tableau comme celui présenté dans le **tableau 20**, qui contient les formules appropriées à utiliser dans Excel.

Tableau 20 : Tableau de distribution des données du modèle (CTT)

	A	B	C	D
1	Score	Fréquence	Pourcentage	Pourcentage cumulatif
2	0	Insérer la fréquence de l'itération la plus récente de l'évaluation	=B2/B(n+1)*100	=C2
3	1	Insérer la fréquence de l'itération la plus récente de l'évaluation	=B3/B(n+1)*100	=D2+C3
4	2	Insérer la fréquence de l'itération la plus récente de l'évaluation	=B4/B(n+1)*100	=D3+C4
...
n	Score maximal	Insérer la fréquence de l'itération la plus récente de l'évaluation	=Bn/B(n+1)*100	=D(n-1)+Cn (cela doit être égal à 100)
n+1	Total	=SUM(B2:Bn)		

Théorie de la réponse aux items

Une approche de notation numérique correcte sera utilisée pour traduire les repères du processus Angoff vers l'emplacement de l'échelle IRT qui sera ensuite utilisé pour calculer le pourcentage d'élèves atteignant chacun des critères de réussite. Pour préparer les distributions, l'analyste de données devra suivre les étapes suivantes :

1. Classer les items utilisés dans l'atelier d'établissement de liens entre les politiques par difficulté sur l'échelle IRT sous-jacente ajustée pour la probabilité de réponse de 0,67.
2. Calculer la proportion d'élèves atteignant et dépassant cette estimation de capacité en utilisant l'approche utilisée lors de la toute dernière itération de l'évaluation.

Ils utiliseront ces informations pour préparer un tableau comme celui présenté dans le **tableau 21**.

Tableau 21 : Tableau de distribution des données du modèle (IRT)

	A	B	C
1	Identifiant de l'item	Difficulté de l'item	Pourcentage estimé atteignant cette difficulté de niveau
2	Item le plus simple	Insérer la difficulté IRT de l'itération la plus récente de l'évaluation	Insérer le pourcentage de l'itération la plus récente de l'évaluation
3	Prochain item le plus simple	Insérer la difficulté IRT de l'itération la plus récente de l'évaluation	Insérer le pourcentage de l'itération la plus récente de l'évaluation
4	Prochain item le plus simple	Insérer la difficulté IRT de l'itération la plus récente de l'évaluation	Insérer le pourcentage de l'itération la plus récente de l'évaluation
...
n	Item le plus difficile	Insérer la difficulté IRT de l'itération la plus récente de l'évaluation	Insérer le pourcentage de l'itération la plus récente de l'évaluation

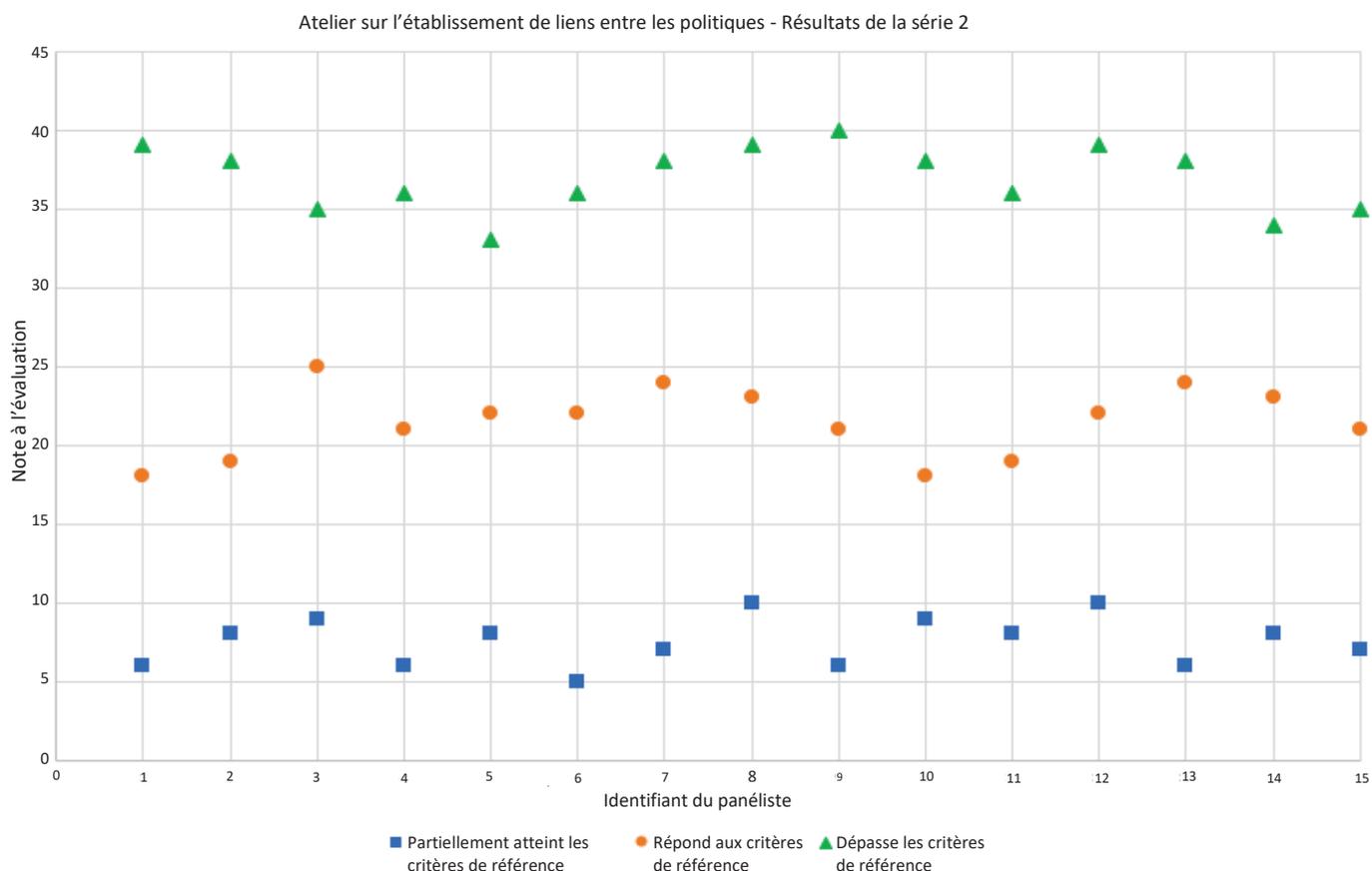
Au cours de l'atelier, les données d'impact pour chaque MPL seront déterminées en utilisant le référentiel du processus Angoff pour identifier lesquels, si les items du **tableau 21**, qui sont classés par ordre de difficulté, devraient être utilisés pour déterminer la proportion d'apprenants atteignant le MPL. Par exemple, si le critère de référence pour compétence minimale « atteinte » est de 13, alors le 13^{ème} item du livret d'items commandés sera utilisé pour déterminer la proportion d'apprenants atteignant la compétence minimale.

ANNEXE O – EXEMPLES DE DONNÉES DE RÉTROACTION ET INSTRUCTIONS

INFORMATIONS NORMATIVES (PARFOIS APPELÉES STATISTIQUES DE LOCALISATION)

Après chaque série de notations, l'analyste de données doit créer un graphique comme celui de la **figure 37** qui montre les numéros uniques de chacun des panélistes (connus d'eux seuls) et leur critère pour chacun des GPL. Le graphique peut être créé à l'aide du type de diagramme Nuage de points dans Excel avec des données sur les critères au niveau des panélistes par GPL.

Figure 37 : exemple de données normatives sur les notes des panélistes



INFORMATIONS SUR LES IMPACTS

Pour générer les informations sur les impacts, l'analyste de données doit prendre les références au niveau du panel fixées par les panélistes pour chaque GPL et, en utilisant les distributions de données, identifier le pourcentage d'apprenants qui tomberaient dans chaque GPL sur la base de l'itération la plus récente de l'évaluation.

Tableau 22 : tableau de données d'impact du modèle

MPL	Critère de référence	Plage de notes	Pourcentage d'apprenants
En dessous du niveau partiellement atteint	N/D	$0 - (x-1)$	Insérer le pourcentage calculé à partir de la répartition des données
Partiellement atteint	x	$x - (y-1)$	Insérer le pourcentage calculé à partir de la répartition des données
Atteint	y	$y - (z-1)$	Insérer le pourcentage calculé à partir de la répartition des données
Dépasse	z	$z - \text{note maximale}$	Insérer le pourcentage calculé à partir de la répartition des données
Total			100.0

ANNEXE P – PROGRAMME DE L’ATELIER

Le **tableau 23** montre les sessions et les horaires de chaque tâche de l’atelier. Certains horaires sont flexibles, car la durée de la session dépendra de facteurs tels que le nombre d’items utilisés et la vitesse à laquelle un consensus est atteint, le cas échéant. L’équipe du projet devra les utiliser pour construire son programme. Toutes les exigences spécifiques à la session sont indiquées dans la colonne « Remarques ». Des exemples de programme pour les ateliers en présentiel et à distance sont présentés à l’**annexe Q** et peuvent être adaptés par les pays si nécessaire.

Tableau 23 : Programme de l’atelier

Tâche	Session	Présentation	Animateur	Durée requise	Remarques
Ouverture	Mots de bienvenue et présentations	1	Animateur principal	30 minutes	Doit être effectué en une seule séance
	Allocution des représentants du ou des gouvernements, de l’agence d’évaluation (le cas échéant) et de l’organisme donateur (le cas échéant)	1		À déterminer par pays	
	Présentation du programme, des objectifs et de la vue d’ensemble méthodologique	2	Animateur principal	30 minutes	
Familiarisation	Familiarisation avec le GPF	3	Tous les animateurs	120 minutes	Peut avoir lieu avant l’atelier
	Familiarisation avec l’outil d’évaluation	4	Animateur de contenu	120 minutes	Peut avoir lieu avant l’atelier
	Évaluation	19	Animateur principal	10 minutes	
Alignement (Tâche 1)	Former les panélistes à la tâche d’alignement (y compris les exercices pratiques)	5	Animateur principal	60 minutes	
	Les panélistes entreprennent une activité d’alignement (activité indépendante)	6	Animateur de contenu	120 minutes	Le temps variera en fonction du nombre d’items. L’activité doit avoir lieu à la fin d’une session pour permettre aux animateurs de disposer de suffisamment de flexibilité et de temps pour rassembler les résultats avant l’activité suivante
	Présentation et discussion sur les résultats de l’alignement	7	Animateur principal	45 minutes	
	Évaluation	19	Animateur principal	10 minutes	
Mise en correspondance (Tâche 2)	Former les panélistes à la tâche de mise en correspondance	8	Animateur principal	60 minutes	
	Les panélistes entreprennent une activité de mise en correspondance (activité de groupe)	9	Animateur de contenu	120 minutes	Le temps variera en fonction du nombre d’items. L’activité doit avoir lieu à la fin d’une session pour permettre une certaine flexibilité, d’autant plus que parvenir à un consensus peut prendre plus de temps
	Présentation et discussion des résultats de la mise en correspondance	10	Animateur principal	30 minutes	
	Évaluation	19	Animateur principal	10 minutes	

Tâche	Session	Présentation	Animateur	Durée requise	Remarques
Analyse comparative (Tâche 3)	Présentation des normes mondiales et de l'approche d'analyse comparative	11	Animateur principal	30 minutes	
	Former les panélistes à la méthode Angoff	12	Animateur principal	60 minutes	
	Les panélistes entreprennent la méthode Angoff avec des exercices pratiques	13	Animateur de contenu	30 minutes	
	Série 1	14	Animateur de contenu	120 minutes	Le temps variera en fonction du nombre d'items. L'activité doit avoir lieu à la fin d'une session pour permettre une certaine flexibilité et donner le temps aux animateurs de rassembler les résultats avant l'activité suivante
	Présentation et discussion des résultats de la série 1 et des données d'impact	15	Animateur principal	60 minutes	
	Présentation de la série 2 Angoff	16	Animateur principal	15 minutes	
	Série 2	17	Animateurs de contenu	120 minutes	Le temps variera en fonction du nombre d'items. L'activité doit avoir lieu à la fin d'une session pour permettre une certaine flexibilité et donner le temps aux animateurs de rassembler les résultats avant l'activité suivante
	Présentation des résultats de la série 2	18	Animateurs principaux	30 minutes	
	Évaluation	19	Animateur principal	10 minutes	
Clôture	Mots de clôture et présentation des certificats	20	À déterminer par pays	À déterminer par pays	

ANNEXE Q – EXEMPLES D’ORDRE DU JOUR POUR UN ATELIER EN PRÉSENTIEL ET À DISTANCE

Tableau 24 : Exemple d’ordre du jour pour un atelier en présentiel

Heure	Activité	Présentation	Animation
09h00-09h30	Inscription		Équipe du projet
09h30-10h45	Ouverture, présentations, ordre du jour et logistique	1	Gouvernement/agence d'évaluation, donateurs et partenaires de mise en œuvre (le cas échéant) ainsi que les principaux animateurs
10h45-11h00	Pause-café		--
11h00-11h45	Contexte, objectif et présentation de la méthode	2	Animateur principal
11h45-12h30	Présentation du GPF et examen des GPD	3	Tous les animateurs
12h30-13h30	Pause déjeuner		--
13h30-14h45	Examen du GPF	3	Tous les animateurs
14h45-15h30	Présentation de l'évaluation/des évaluations	4	Animateur de contenu
15h30-15h45	Pause-café		--
15h45-16h45	Examen des items de l'évaluation	4	Animateur de contenu
16h45-17h00	Clôture du jour 1 et aperçu du jour 2		Animateur principal
Heure	Jour 2	Présentation	Animation
09h00-09h30	Présentation du jour 2 et résolution des problèmes du jour 1		Animateur principal
09h30-10h30	Passer l'évaluation		Animateur de contenu
10h30-10h45	Pause-café		--
10h45-12h30	Examen du GPF et identification des problèmes restants		Animateur principal
12h30-13h30	Pause déjeuner		--
13h30-14h15	Discussion visant à clarifier les questions en suspens de la séance du matin		Tous les animateurs
14h15-15h00	Tâche 1 : formation sur l'exercice d'alignement	5	Animateur principal
15h00-15h15	Pause-café		--
15h15-16h00	Tâche 1 : discussions en petits groupes sur les 5 premiers items	6	Animateur de contenu
16h00-16h45	Tâche 1 : discussion en plénière sur les questions soulevées en groupe	6	Animateur de contenu
16h45-17h00	Clôture du jour 2 et aperçu du jour 3		Animateur principal
Heure	Jour 3	Présentation	Animation
09h00-09h15	Présentation du jour 3 et résolution des problèmes du jour 2		Animateur principal
09h15-10h30	Tâche 1 : exercice d'alignement	6	Animateur de contenu
10h30-10h45	Pause-café		--
10h45-12h30	Tâche 1 : exercice d'alignement (suite)	6	Animateur de contenu
12h30-13h30	Pause déjeuner		--
13h30-14h15	Tâche 1 : présentation et discussion des résultats de l'alignement	7	Animateur principal
14h15-15h00	Tâche 2 : formation à l'exercice de mise en correspondance	8	Animateur de contenu
15h00-15h15	Pause-café		--
15h15-16h45	Tâche 2 : début de l'exercice de mise en correspondance	9	Animateur de contenu
16h45-17h00	Clôture du Jour 3 et aperçu du Jour 4		

Heure	Activité	Présentation	Animation
Heure	Jour 4	Présentation	Animation
09h00-09h15	Présentation du jour 4 et résolution des problèmes du jour 3		Animateur principal
09h15-10h30	Tâche 2 : des petits groupes effectuent ensemble l'exercice de mise en correspondance	9	Animateur de contenu
10h30-10h45	Pause-café		--
10h45-12h30	Tâche 2 : discussion en plénière sur les résultats de la mise en correspondance	10	Animateur de contenu
12h30-13h30	Pause déjeuner		--
13h30-14h00	Présentation des normes mondiales et de l'approche d'analyse comparative	11	Animateur principal
14h00-14h45	Tâche 3 : formation sur la méthode Angoff	12	Animateur principal
14h45-15h15	Tâche 3 : pratique Angoff en petits groupes	13	Animateur principal
15h15-15h30	Pause-café		--
15h30-16h15	Tâche 3 : Discussion en plénière des questions soulevées en petits groupes		Animateur principal
16h15-16h45	Tâche 3 : Série 1 Angoff	14	Animateur de contenu
16h45-17h00	Clôture du jour 4 et aperçu du jour 5		Animateur principal
17h00-18h00	Heure de consultation au cours de laquelle les panélistes peuvent consulter les animateurs		Tous les animateurs
Heure	Jour 5	Présentation	Animation
09h00-09h15	Présentation du jour 5 et résolution des problèmes du jour 4		Animateur principal
09h15-10h30	Tâche 3 : Poursuivre les évaluations de la série 1	14	Animateur de contenu
10h30-10h45	Pause-café		--
10h45-12h30	Tâche 3 : Terminer les évaluations de la série 1	14	Animateur de contenu
12h30-13h30	Pause déjeuner		--
13h30-15h15	Tâche 3 : Présentation et discussion des résultats de la série 1	15	Animateur principal
15h15-15h30	Pause-café		--
15h30-16h30	Tâche 3 : Examiner les notes de la première série en petits groupes pour discuter de tous les items sur lesquels un désaccord s'est exprimé		Animateur de contenu
16h30-17h00	Tâche 3 : Partager et discuter des données sur la difficulté des items et leur impact	15	Animateur principal
17h00-17h15	Clôture du jour 5 et aperçu du jour 6		Animateur principal
17h15-18h15	Heure de consultation au cours de laquelle les panélistes peuvent consulter les animateurs		Tous les animateurs
Heure	Jour 5	Présentation	Animation
09h00-09h15	Présentation du jour 6 et résolution des problèmes du jour 5		Animateur principal
09h15-09h30	Tâche 3 : Présentation de la série 2 Angoff	16	Animateur principal
09h30-10h30	Tâche 3 : Série 2 Angoff	17	Animateur de contenu
10h30-10h45	Pause-café		--
10h45-12h30	Tâche 3 : terminer les évaluations de la série 2	17	Animateur de contenu
12h30-13h30	Pause déjeuner		--
13h30-14h30	Évaluation de l'atelier	19	Animateur principal
14h30-15h15	Tâche 3 : Présentation des résultats de la série 2	18	Animateur principal
15h15-15h30	Pause-café		--
15h30-16h30	Discuter des résultats et des dernières questions des panélistes		Animateur principal
16h30-17h00	Clôture et logistique	20	Animateur principal

Tableau 25 : Exemple d'ordre du jour pour la session de préparation à distance 1*(Il est recommandé qu'elle ait lieu deux semaines avant l'atelier)*

Durée	Activité	Présentation	Animateur
0–15 minutes	Mots de bienvenue et présentations	1	Animateur principal
15–40 minutes	Présentation de l'établissement de liens entre les politiques	2	Animateur principal
40–55 minutes	Objectif de la séance de préparation		Animateur du processus
55–60 minutes	Courte pause		
60–80 minutes	Présentation du GPF	3	Animateur de contenu ou principal
80–100 minutes	[Niveau et matière] Examen du GPF	3	Animateur de contenu ou principal
100–110 minutes	Explication des activités entre les sessions		Animateur principal
110–120 minutes	Conclusion		Animateur principal

Activités entre les sessions des panélistes :

- Examiner [niveau et sujet] le GPF et identifier tous les items qui ne sont pas clairs (soumettre une semaine avant l'atelier)

Tableau 26 : Exemple d'ordre du jour pour la session de préparation à distance 2*(Il est recommandé qu'elle ait lieu deux jours après la première séance préparatoire)*

Durée	Activité	Présentation	Animateur
0–15 minutes	Mot de bienvenue et objectif de la séance de préparation		Animateur principal
15–30 minutes	Présentation du [nom de l'évaluation]	4	Animateur de contenu ou animateur principal
30–55 minutes	Examiner chaque item de [l'évaluation]	4	Animateur de contenu ou animateur principal
55–60 minutes	Courte pause		
60–100 minutes	Continuer à examiner les items et discuter de l'administration de [l'évaluation]	4	Animateur de contenu ou animateur principal
100–110 minutes	Explication des activités entre les sessions		Animateur principal
110–120 minutes	Conclusion		Animateur principal

Activités entre les sessions des panélistes :

- Administrer [l'évaluation] à trois apprenants (du niveau/groupe d'âge approprié pour chaque GPL)

Tableau 27 : Exemple d'ordre du jour pour la session 1 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mot de bienvenue et objectif de la session 1		Animateur principal
10–55 minutes	Examiner l'activité du GPF et apporter des éclaircissements		Animateur de contenu ou animateur principal
55–60 minutes	Courte pause		
60–105 minutes	Discussion sur l'activité d'administration [de l'évaluation]		Animateur de contenu ou animateur principal
105–120 minutes	Approche d'évaluation et réalisation de l'évaluation 1	19	Animateur principal

Tableau 28 : Exemple d'ordre du jour pour la session 2 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mot de bienvenue et objectif de la séance 2		Animateur principal
10–20 minutes	Répondre à toutes les préoccupations soulevées lors de l'évaluation 1		Animateur de contenu ou animateur principal
20–55 minutes	Introduction à la tâche d'alignement (tâche 1)	5	Animateur principal
55–60 minutes	Courte pause		
60–90 minutes	Discussions en petits groupes sur les 5 premiers items ²⁵	6	Animateurs de contenu ²¹
90–110 minutes	Discussion en plénière sur les questions soulevées dans les groupes	6	Animateur principal
110–120 minutes	Explication des activités entre les sessions et clôture		Animateur principal

Activités entre les sessions des panélistes :

- Terminer la tâche 1 : examen de l'alignement sur tous les items restants (soumettre quatre heures après la séance)
- Terminer l'évaluation 2 (soumettre avec la révision de l'alignement)

Tableau 29 : exemple de programme pour la session 3 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mots de bienvenue et objectif de la session 3		Animateur principal
10–40 minutes	Examiner les activités entre les sessions et apporter des éclaircissements	7	Animateur de contenu
40–55 minutes	Introduction à la tâche 2 – Mise en correspondance avec les GPL et les GPD	8	Animateur principal
55–120 minutes	Pratiquer avec la tâche 2		Animateur principal
120–130 minutes	Courte pause		
130–230 minutes	Les petits groupes effectuent la tâche 2 ensemble (groupes organisés par niveau/matière/langue) ²⁶	9 & 10	Animateur de contenu
230–240 minutes	Explication des activités entre les sessions et clôture		Animateur principal

Activités entre les sessions des panélistes :

- Compléter l'évaluation 3 (à soumettre une heure après la clôture de la session)

Tableau 30 : Exemple d'ordre du jour pour la session 4 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mots de bienvenue et objectif de la session 4		Animateur principal
10–40 minutes	Présenter la méthodologie Angoff et la tâche 4 puis apporter des éclaircissements	11 & 12	Animateur principal
40–75 minutes	Évaluations Angoff en petits groupes à l'aide d'exercices pratiques	13	Animateur de contenu ou animateur principal
75–80 minutes	Courte pause		
80–100 minutes	Discussion en plénière des questions soulevées en petits groupes		Animateur principal
100–110 minutes	Commencer les évaluations de la série 1 (soulever les questions qui surviennent)	14	Travail autonome
110–120 minutes	Explication des activités entre les sessions et clôture		Animateur principal

²⁵ Chaque petit groupe aura un animateur de contenu ; nous recommandons au(x) animateur(s) principal(aux) de rester en dehors des petits groupes afin que les petits groupes puissent identifier leurs questions et les rapporter en plénière.

²⁶ Ibid.

Activités entre les sessions des panélistes :

- Réunions individuelles entre chaque panéliste et un animateur principal (au cours de ces réunions, les animateurs répondent aux questions des panélistes et leur demanderont comment ils notent chaque item et pourquoi et vérifieront que le raisonnement suit le déroulement des étapes requises pour cela tâche)
- Compléter les évaluations de la série 1 sur tous les items restants (soumettre quatre heures après la clôture de la session ou une heure après la réunion individuelle avec les principaux animateurs, selon la dernière éventualité)
- Terminer l'évaluation 4 (soumettre avec les notations de la série 1)

Tableau 31 : Exemple d'ordre du jour pour la session 5 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mots de bienvenue et objectif de la session 5		Animateur principal
10–45 minutes	Examiner et discuter des notations de la série 1 en plénière	15	Animateur de contenu
45–50 minutes	Courte pause		
50–110 minutes	Examiner en petits groupes (organisés par niveau/matière/langue) les notations de la série 1, en passant en revue chaque item pour lequel un désaccord est survenu	15	Animateur de contenu
110–150 minutes	Partager et discuter des données sur la difficulté des items et leur impact	15	Animateur principal
150–180 minutes	Explication des activités entre les sessions (rappel de la méthodologie) et clôture	16	Animateur principal

Activités entre les sessions des panélistes :

- Réunions individuelles entre chaque panéliste et un animateur principal (au cours de ces réunions, les animateurs répondent aux questions des panélistes et leur demanderont comment ils notent chaque item et pourquoi et vérifieront que le raisonnement suit le déroulement des étapes requises pour cela tâche)
- Compléter les évaluations de la série 2 (soumettre quatre heures après la clôture de la session ou une heure après la réunion individuelle avec les principaux animateurs, selon la dernière éventualité)
- Effectuer l'évaluation 5

Tableau 32 : Exemple d'ordre du jour pour la session 6 de l'atelier à distance

Durée	Activité	Présentation	Animateur
0–10 minutes	Mots de bienvenue et objectif de la session 6		Animateur principal
10–30 minutes	Examiner les notes de la série 2 et partager les résultats finaux	18	Animateur de contenu
30–90 minutes	Discuter des résultats et des dernières questions des panélistes	18	Animateur principal
90–100 minutes	Effectuer l'évaluation 6	19	Travail autonome
100–120 minutes	Remerciements et clôture	20	Animateur principal

ANNEXE R – FORMULAIRE D'ÉVALUATION DE L'ATELIER

Ce formulaire peut être soit découpé de manière à ce que chacune des sections soit administrée après la journée/session de l'atelier au cours de laquelle le sujet est présenté, soit administré dans son intégralité le dernier jour/séance de l'atelier. La gestion de l'atelier tout du long aidera les animateurs à identifier les lacunes dans la compréhension et à adapter leurs présentations si nécessaire, mais cela peut également être trop fastidieux pour les panélistes. Les animateurs doivent prendre une décision en consultation avec les principales parties prenantes en fonction du contexte de l'atelier. Le langage d'introduction de chaque section doit être adapté en fonction du moment où les questions sont présentées. Vous devrez également remplir tout ce qui est entre parenthèses. Enfin, certaines questions devront peut-être être déplacées vers une autre session pour les ateliers à distance où les activités n'ont pas toujours lieu le même jour que la formation. Quelle que soit la manière dont le formulaire est présenté, il est important d'inclure l'identifiant du panéliste sur l'ensemble du formulaire (s'il est administré dans un seul contexte) ou au moins pour les évaluations des séries 1 et 2 (s'il est administré au cours de l'atelier).

PARTIE 1 : FORMATION SUR LE CADRE MONDIAL DE COMPÉTENCES

Aujourd'hui, vous avez été formé sur les descripteurs mondiaux de compétences (GPDs). Veuillez lire attentivement les affirmations suivantes et cochez cette catégorie pour indiquer votre niveau d'accord.

Tableau 33 : Formulaire d'évaluation de la formation sur le GPF

Formation sur les GPD	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le but du GPF					
Je comprends la relation entre les domaines, les concepts, les sous-concepts, les connaissances et les compétences, et les GPD					
Les GPD étaient clairs et faciles à comprendre					
La discussion sur les GPD m'a aidé à comprendre ce que l'on attend des apprenants de [insérer la matière] à la fin de [insérer la note]					
L'exercice pratique utilisant les GPD m'a permis d'améliorer ma compréhension					
Chacun a pu partager ses idées et opinions de manière égale					
Chacun a pu poser ses questions de manière égale					
Le temps consacré à la formation sur le GPD était suffisant					

Veuillez décrire avec vos propres mots quel est l'objectif du GPF et ce que vous savez des GPD.

Veuillez énumérer toutes les questions ou zones de confusion que vous avez au sujet du GPF.

Veuillez énumérer tous les conseils/demandes destinés aux animateurs qui pourraient rendre la formation plus efficace pour vous.

PARTIE II : FORMATION SUR L'ÉVALUATION /LES ÉVALUATION(S)

Aujourd'hui, vous avez été formé sur l'évaluation ou les évaluations que nous utiliserons pour établir des liens entre les politiques. Veuillez lire attentivement les affirmations suivantes et cochez chaque catégorie pour indiquer dans quelle mesure vous êtes d'accord avec chaque d'elles.

Tableau 34 : Formulaire d'évaluation pour la formation à l'évaluation

Formation à l'évaluation	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le but de l'atelier					
Je comprends les concepts évalués dans l'évaluation					
Je comprends comment l'évaluation est administrée					
L'administration de l'évaluation m'a aidé à comprendre les performances des apprenants peu compétents à l'évaluation (cela n'est applicable que si les panélistes ont pu évaluer les apprenants avant l'atelier)					
Je pense avoir une bonne idée des résultats obtenus par les apprenants peu compétents à l'évaluation					
Le temps consacré à la formation à l'évaluation était suffisant					

Veuillez énumérer toutes les questions que vous avez sur l'évaluation ou les évaluations.

Veuillez énumérer tous les conseils/demandes destinés aux animateurs qui pourraient rendre la formation plus efficace pour vous.

PARTIE III : FORMATION À LA MÉTHODE D'ALIGNEMENT

Aujourd'hui, vous avez été formé à la méthodologie d'alignement. Veuillez lire attentivement les affirmations suivantes et cocher chaque catégorie pour indiquer dans quelle mesure vous êtes d'accord avec chacune d'elles.

Tableau 35 : Formulaire d'évaluation pour la tâche 1 – Alignement

Formation sur l'alignement	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le but de l'alignement					
Je comprends la méthodologie d'alignement					
Je comprends la différence entre aucune correspondance, correspondance partielle et correspondance complète					
Je suis confiant quant à mes notes d'alignement					
Le temps consacré à la formation à l'évaluation était suffisant					

Veuillez énumérer toutes vos questions ou zones de confusion concernant la méthodologie/le processus d'alignement.

Veillez énumérer tous les conseils/demandes destinés aux animateurs qui pourraient rendre la formation plus efficace pour vous.

PARTIE IV : FORMATION À LA MÉTHODOLOGIE DE LA MISE EN CORRESPONDANCE

Aujourd'hui, vous avez été formé à la méthodologie de la mise en correspondance. Veuillez lire attentivement les affirmation suivantes et cocher chaque catégorie pour indiquer dans quelle mesure vous êtes d'accord avec chacune d'elles.

Tableau 36 : formulaire d'évaluation pour la tâche 2 – Mise en correspondance

Formation sur l'alignement	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le but de la mise en correspondance					
Je comprends la méthodologie de mise en correspondance					
Je comprends comment l'activité d'alignement est liée à l'activité de mise en correspondance					
Je suis d'accord avec le consensus du groupe sur les GPL et GPD sur lesquels nous avons aligné chaque item (développez ci-dessous sinon)					
Le temps consacré à la formation à la mise en correspondance était suffisant					

Veillez décrire toute décision de groupe concernant la mise en correspondance avec laquelle vous n'êtes pas d'accord et la raison.

Veillez énumérer toutes les questions ou zones de confusion que vous avez concernant la méthodologie/le processus de mise en correspondance.

Veillez énumérer tous les conseils/demandes destinés aux animateurs qui pourraient rendre la formation plus efficace pour vous.

PARTIE V : FORMATION SUR LA MÉTHODOLOGIE DE L'ANALYSE COMPARATIVE (ANGOFF)

Aujourd'hui, vous avez été formé à la méthodologie de l'analyse comparative. Veuillez lire attentivement les affirmation suivantes et cocher chaque catégorie pour indiquer dans quelle mesure vous êtes d'accord avec chacune d'elles.

Tableau 37 : Formulaire d'évaluation pour la tâche 3 – Analyse comparative

Formation sur l'établissement de liens entre les politiques	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le processus à suivre pour réaliser l'exercice d'analyse comparative					
Je comprends comment la méthodologie d'analyse comparative est liée aux étapes d'alignement et de mise en correspondance					

Formation sur l'établissement de liens entre les politiques	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends le niveau de difficulté des items d'évaluation					
La discussion sur la procédure a été suffisante pour me permettre d'avoir confiance en la méthodologie					
Je comprends comment mes notes donneront lieu à un critère de					
Chacun a pu partager ses idées et opinions de manière égale					
Chacun a pu poser ses questions de manière égale					
Le temps consacré à la formation sur la méthode d'établissement de liens entre les politiques était suffisant					
Je suis confiant quant à mes notes de la série 1					

Veillez décrire la méthodologie d'analyse comparative avec vos propres mots.

Veillez énumérer toutes vos questions ou zones de confusion concernant la méthodologie/le processus d'analyse comparative.

Veillez énumérer tous les conseils/demandes destinés aux animateurs qui pourraient rendre la formation plus efficace pour vous.

PARTIE VI : ÉVALUATION DE L'ANALYSE COMPARATIVE DE LA SÉRIE 2

Au cours de la série 2, vous avez reçu des informations sur les performances réelles et des données sur l'impact de l'utilisation des résultats de la série 1. Ensuite, il vous a été demandé de donner des prévisions de performances révisées. Veuillez sélectionner la meilleure réponse ci-dessous.

Tableau 38 : Formulaire d'évaluation pour la tâche 3 – Série 2 d'analyse comparative

Série 2	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
Je comprends les données sur les notes des autres					
Je comprends les données de difficulté de l'item et leur rapport avec ce processus					
Je comprends les données d'impact et leur rapport avec ce processus					
Je suis confiant quant aux prévisions de performances que j'ai faites lors de la série 2					
Mes prédictions de performances ont été influencées par les informations montrant les notes des autres panélistes					
Mes prédictions de performances ont été influencées par les données de difficulté des items en montrant les performances réelles des apprenants lors de l'évaluation					
Mes prédictions de performances ont été influencées par les informations d'impact montrant les résultats pour l'échantillon d'apprenants					

Série 2	En total désaccord	En désaccord	Neutre	D'accord	Tout à fait d'accord
J'ai eu suffisamment de temps pour effectuer les prévisions de performances de la série 2					

Avez-vous des commentaires supplémentaires sur la série 2 ?

Partie V : Évaluation globale

Dans quelle mesure êtes-vous à l'aise avec vos prédictions de performances finales ?

Très mal à l'aise	Quelque peu mal à l'aise	Plutôt à l'aise	Très à l'aise

Si vous avez coché l'une des options indiquant un malaise, veuillez expliquer la raison.

Dans l'ensemble, comment évaluez-vous la réussite de l'atelier consacré à l'établissement de liens entre les politiques ?

- a. Très satisfaisant
- b. Satisfaisant
- c. Insatisfaisant
- d. Très insatisfaisant

Comment jugez-vous l'organisation de l'atelier ?

- a. Très satisfaisant
- b. Satisfaisant
- c. Insatisfaisant
- d. Très insatisfaisant

Veillez fournir tout commentaire qui, selon vous, pourrait nous être utile dans la planification des futurs ateliers consacrés à l'établissement de liens entre les politiques.

Nous vous remercions pour votre participation à l'atelier.

ANNEXE S – DIAPOSITIVES À L'INTENTION DES ANIMATEURS DE CONTENU

Il est essentiel que tous les animateurs soient formés à la méthodologie d'établissement de liens entre les politiques. Toutefois, d'une manière générale, les animateurs principaux auront été formés avant le processus d'établissement de liens entre les politiques. Il est donc probable que seuls les animateurs de contenu devront être formés. Les animateurs principaux doivent obtenir le contenu des diapositives de formation des animateurs à partir des diapositives de l'atelier incluses à **l'annexe G**. Nous recommandons au moins huit heures de formation pour les animateurs de contenu avant l'atelier, bien que cela puisse varier en fonction de leur expérience de l'établissement de normes en général, de l'évaluation et de la méthode Angoff modifiée. Les diapositives doivent être réduites pour laisser le temps de parcourir tout le contenu technique majeur, en mettant l'accent sur les présentations suivantes :

- Présentation 3
- Présentation 5
- Présentation 7
- Présentation 8
- Présentation 10
- Présentation 11
- Présentation 12
- Présentation 15
- Présentation 18.

Il est particulièrement essentiel que les animateurs de contenu aient une compréhension approfondie du GPF et de l'évaluation, car comprendre et relayer ce contenu et le replacer dans le contexte local est leur principale responsabilité. Il est utile que les animateurs de contenu comprennent également à quel moment les différents sujets/vocabulaire/etc. sont enseignés dans les écoles dans le contexte local, quelle terminologie est utilisée en classe, etc.

La formation doit également couvrir les choses à faire et à ne pas faire lors de l'animation de l'atelier fourni dans le **tableau 11**.

Si le temps entre la formation et l'atelier est suffisant, il peut être utile d'entreprendre une répétition des sections pertinentes de l'atelier avec les animateurs de contenu, les animateurs principaux agissant en tant que panélistes, pour garantir la compréhension.

ANNEXE T – CALCULS DES CRITÈRES DE RÉFÉRENCE POUR L'ATELIER

CALCUL DES CRITÈRES DE RÉFÉRENCE POUR LA MÉTHODE ANGOFF

Les critères pour atteindre partiellement, atteindre et dépasser la compétence minimale sont calculés à l'aide d'un ensemble de six équations. Les équations un à trois sont utilisées pour calculer les critères de référence pour chaque panéliste et les équations quatre à six sont utilisées pour calculer les critères de référence recommandés par le panel. Pour ces équations, i indique les items ou les mots, j indique les panélistes, l indique le nombre d'items ou de mots tentés par JPA, m indique le nombre d'items ou de mots tentés par JA et n indique le nombre d'items ou de mots tentés par JD. Lorsqu'on ne fixe qu'un seul critère de référence, au lieu de trois, le calcul est beaucoup plus simple. Dans ce cas, il vous suffit d'additionner le total des oui pour les réponses des panélistes, puis de faire la moyenne de ces totaux entre les panélistes.

L'équation 1 montre le critère de référence Partiellement atteint les compétences minimales pour un panéliste à l'issue de la série I.

$$PM_j = \sum_{i=1}^l JP_{ij} \quad (1)$$

L'équation 2 montre le critère de référence Atteint les compétences minimales pour un panéliste à l'issue de la série I.

$$M_j = PM_j + \sum_{i=l+1}^m JM_{ij} \quad (2)$$

L'équation 3 montre le critère de référence Dépasse les compétences minimales pour un panéliste à l'issue de la série I.

$$E_j = M_j + \sum_{i=m+1}^n JE_{ij} \quad (3)$$

L'équation 4 montre le critère de référence Partiellement atteint les compétences minimales pour tous les panélistes après la série I.

$$P = \frac{1}{z} \sum_{j=1}^z \sum_{i=1}^l PM_{ij}$$

L'équation 5 montre le critère de référence Atteint les compétences minimales tous les panélistes après la série I.

$$M = \frac{1}{z} \sum_{j=1}^z (PM_j + \sum_{i=l+1}^m M_{ij}) \quad (5)$$

L'équation 6 montre le critère de référence Dépasse les compétences minimales pour tous les panélistes après la série I.

$$E = \frac{1}{z} \sum_{j=1}^z (M_j + \sum_{i=m+1}^n E_{ij}) \quad (6)$$

Comment calculer les critères de référence

Étape 1 : Calculer le score pour Partiellement atteint les compétences minimales (PA_i) pour un panéliste après la série I.

- i. Déterminer combien d'items ou de mots le panéliste a décidé que deux des trois apprenants de niveau atteint les compétences minimales peuvent tenter de répondre ou de lire en une minute (applicable uniquement pour une tâche chronométrée).
- ii. En considérant uniquement ces items ou mots auxquels deux des trois apprenants de niveau juste partiellement atteint les compétences minimales (JPA), les apprenants peuvent répondre ou lire correctement selon le panéliste, additionner tous les items ou mots de ce sous-ensemble que le panéliste a évalué comme atteignant seulement partiellement les compétences minimales.
- iii. PA_j pour ce panéliste est la somme de la sous-étape 2
- iv. Répéter les sous-étapes 1 et 2 pour chaque panéliste afin de calculer PA_j pour chacun

Étape 2 : Calculer le score Atteint les compétences minimales (A_j) pour un panéliste après la série I.

- i. Déterminer le nombre d'items ou de mots que le panéliste a décidé que deux des trois apprenants de niveau juste atteint les compétences minimales peuvent tenter de répondre ou de lire en une minute (applicable uniquement pour la tâche chronométrée).
- ii. En considérant uniquement les items ou les mots auxquels deux des trois apprenants de niveau juste atteint les compétences minimales peuvent répondre ou lire correctement selon le panéliste, additionner tous les items de ce sous-ensemble que le panéliste a évalué comme atteignant partiellement et atteignant à peine les compétences minimales.
- iii. A_j pour ce panéliste est la somme de la sous-étape 2.
- iv. Répéter les sous-étapes 1 et 2 pour chaque panéliste afin de calculer A_j pour chacun.

Étape 3 : calculer le score de Dépasse les compétences minimales (D_j) pour un panéliste après la série I.

- i. Déterminer combien d'items ou de mots le panéliste a décidé que deux des trois apprenants de niveau juste dépasse les compétences minimales peuvent tenter de répondre ou de lire en une minute (applicable uniquement pour une tâche chronométrée).
- ii. En considérant uniquement les items ou les mots auxquels deux des trois apprenants de niveau Dépasse à peine les compétences minimales peuvent répondre ou lire correctement selon le panéliste, additionner tous les items de ce sous-ensemble que le panéliste a évalué comme juste partiellement atteint, juste atteint et juste dépasse les compétences minimales.
- iii. D_j pour ce panéliste est la somme de la sous-étape 2.
- iv. Répéter les sous-étapes 1 et 2 pour chaque panéliste afin de calculer D_j pour chacun.

Étape 4 : calculer le score de coupure du niveau partiellement atteint les compétences minimales (P) pour tous les panélistes après la série I.

- i. Additionner tous les scores de coupure PA_j des panélistes.
- ii. Diviser la somme des scores de coupure PA_j et la diviser par le nombre total de panélistes.
- iii. Ce résultat est une moyenne simple équivalente à P .

Étape 5 : calculer le score de coupure du niveau atteint les compétences minimales (A) pour tous les panélistes après la série I.

- i. Additionner tous les scores de coupure A_j des panélistes.
- ii. Diviser la somme des scores de coupure A_j et la diviser par le nombre total de panélistes.
- iii. Ce résultat est une moyenne simple équivalente à A .

Étape 6 : calculer le score de coupure du niveau dépasse les compétences minimales (D) pour tous les panélistes après la série I.

- i. Additionner tous les scores de coupure D_j des panélistes.
- ii. Diviser la somme des scores de coupure D_j et la diviser par le nombre total de panélistes.
- iii. Ce résultat est une moyenne simple équivalente à D .

ANNEXE U – MODÈLE DE CERTIFICAT D'APPRÉCIATION

[Insérer le logo du gouvernement et des partenaires en charge du développement]

CERTIFICAT DE PARTICIPATION

Il s'agit de certifier que

[Insérer le nom]

a participé avec succès à **[Insérer le nom du pays]** l'ATELIER CONSACRÉ À L'ÉTABLISSEMENT DE LIENS ENTRE LES POLITIQUES POUR LA PRODUCTION ET LE RAPPORT SUR L'ODD

4.1 pour le niveau **[Insérer le niveau]** le **[Insérer la date, le Mois, l'Année]** organisé par

[Insérer le nom de l'agence gouvernementale organisatrice] en partenariat avec **[Insérer le nom du partenaire de développement]**

à **[Insérer le lieu]**.

Nom

Fonction
Nom du DP

Nom

Fonction
Gouvernement, organisateur

DATE

ANNEXE V – RÉSUMÉ DU MODÈLE D'AUTO-ÉVALUATION (RÉSULTATS DE L'ATELIER)

Outil d'évaluation	[Insérer le nom de l'outil]
Juridiction	[Insérer la juridiction dans laquelle l'outil d'évaluation est administré]
Niveau	[Insérer la note évaluée par l'outil]
Niveau ODD 4.1.1	a / b / c [rayer la mention inutile]
Sujet	Mathématiques / Lecture [rayer la mention inutile]
MPL en cours de définition	Partiellement atteint / Atteint / Dépasse [rayer la mention inutile]
Date de l'atelier d'établissement de liens entre les politiques	[Insérer la date à laquelle l'atelier a eu lieu]
Évaluateurs	[Insérer les noms et les organisations des évaluateurs]

Critère 1 – Tous les panélistes ont-ils rempli les conditions de participation ?	Oui / Non [rayer la mention inutile]
Critère 2 – Le groupe de panélistes était-il suffisamment représentatif au regard des caractéristiques convenues par le pays ?	Oui / Non [rayer la mention inutile]
Critère 3 : Toutes les valeurs aberrantes ont-elles été supprimées avant le calcul des critères de référence finaux ?	Oui / Non [rayer la mention inutile]
Critère 4 – Les critères de référence n'ont-ils été définies que pour les MPLS qui ne présentent pas d'effets de plancher ou de plafond ?	Oui / Non [rayer la mention inutile]
Critère 5 – La statistique de cohérence entre évaluateurs est-elle supérieure ou égale à 0,7 ?	Oui / Non [rayer la mention inutile]
Critère 6 – L'erreur type pour chaque critère de référence a-t-elle été calculée et examinée pour être déterminée le cas échéant ?	Oui / Non [rayer la mention inutile]
Critère 7 – L'intervalle de confiance pour chaque critère de référence a-t-il été calculé et revu pour être déterminé le cas échéant ?	Oui / Non [rayer la mention inutile]
Critère 8 – La note minimale pour chaque section de l'évaluation était-elle supérieure ou égale à 4 ?	Oui / Non [rayer la mention inutile]
Critère 9 – La moyenne des notes de l'évaluation globale était-elle supérieure ou égale à 3 ?	Oui / Non [rayer la mention inutile]

Note d'auto-évaluation globale

L'atelier consacré à l'établissement de liens entre politiques a-t-il satisfait aux 10 critères d'auto-évaluation ?	Oui / Non [rayer la mention inutile]
---	--------------------------------------

Rapport sur l'atelier consacré à l'établissement de liens entre les politiques

Outre le résumé de l'auto-évaluation, l'équipe du projet peut souhaiter produire un rapport sur les résultats de l'atelier consacré à l'établissement de liens entre les politiques. Les rubriques suivantes peuvent être utiles pour l'élaboration d'un tel rapport.

1. Résumé analytique
2. Aperçu de l'évaluation
 - a. Introduction
 - b. Objectif de l'évaluation
 - c. Conception de l'évaluation
 - d. Administration d'échantillonnage et de test
 - e. Notation
3. Résultats de l'auto-évaluation
 - a. Critère 1 : Alignement
 - b. Critère 2 : Examen de l'item
 - c. Critère 3 : Échantillon

- d. Critère 4 : Administration
- e. Critère 5 : Fiabilité
- 4. Méthodologie d'établissement de liens entre les politiques
 - a. Sélection et description des panélistes
 - b. Procédure
 - i. Préparation d'un atelier d'établissement de liens entre les politiques
 - ii. Conduire un atelier d'établissement de liens entre les politiques
 - iii. Finaliser les MPL
 - c. Analyse des notations des séries 1 et 2
- 5. Résultats de l'établissement de liens entre les politiques
 - a. Résultats de la série 1
 - b. Données de rétroaction
 - c. Résultats de la série 2
- 6. Évaluation du processus d'établissement de liens entre les politiques
 - a. Évaluation procédurale (séries 1 et 2)
 - b. Erreur type d'évaluation interne de la moyenne (séries 1 et 2), de la cohérence entre les panélistes (série 2) et des coefficients de concordance et de cohérence (série 2)
- 7. Synthèse des résultats de l'auto-évaluation de l'atelier
- 8. Conclusions et recommandations
- 9. Références
- 10. Annexes
 - a. Liste de contrôle pour la sélection de méthode
 - b. Formulaire de notation
 - c. Formulaire d'évaluation
 - d. Répartition des fréquences du score du test de l'apprenant
 - e. Niveau de difficulté des items du test
 - f. Autres documents et données pertinents