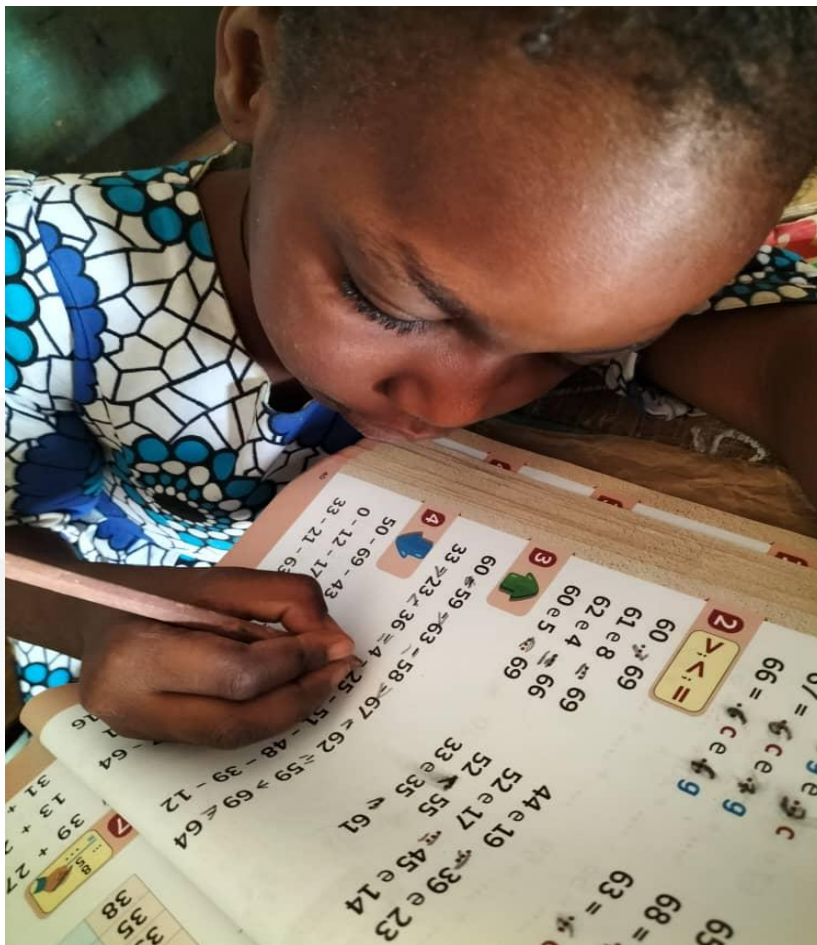


Évaluation du programme de rattrapage scolaire rattrapage scolaire *Ndaw Wune*

Rapport final

Institut Américain de Recherche, avec l'appui de la Fondation Gates

MARS 2026



Sommaire

Remerciements	v
Résumé.....	1
1. Introduction.....	15
2. Théorie du changement	17
3. Contexte	19
4. Questions de recherche	21
5. Méthodes	22
5.1.Évaluation d'impact.....	22
5.2.Analyse des données de suivi.....	33
5.3. Analyse des coûts.....	33
5.4. Analyse coût-efficacité.....	38
6. Résultats	39
6.1.Évaluation de l'impact.....	39
6.2.Analyse des données de suivi.....	51
6.3. Analyse des coûts.....	52
6.4. Analyse coût-efficacité.....	59
7. Conclusions.....	60
8. Recommandations pour accroître davantage l'impact.....	63
Références	68
Annexe A. Test de sélection d'échantillons	A-1
Annexe B. Lieux de mise en œuvre du programme.....	B-1
Annexe C. Méthodologie détaillée de calcul des coûts	C-1
Annexe D. Tests de robustesse	D-1
Annexe E. Fiabilité inter-évaluateurs.....	E-1
Annexe F. Instruments d'enquête	F-1
Questions préalables à l'enquête	F-1
Informations à caractère général	F-2
Évaluation de la lecture	F-6
Évaluation mathématique	F-19

Annexes

1 . Approche méthodologique	4
2 . Aperçu du programme <i>Ndaw Wune</i> en 2024-2025.....	16
3 . Théorie du changement pour <i>Ndaw Wune Année Scolaire</i>	18
Figure4 . Approche DID	23
5 . Emplacement des écoles de traitement et de comparaison	24
Tableau 6. Équilibre au niveau des écoles (variables appariées)	24
Tableau 7. Équivalence de base au niveau scolaire (variables collectées au début de l'étude).....	25
8 . Taille des échantillons d'élèves.....	28
Tableau 9 . Évaluation des compétences en lecture-écriture et mathématiques	29
Tableau 10 . Caractéristiques de base au niveau des élèves.....	39
11	41
Figure12 . Langue dominante chez les élèves	42
13 . Adéquation entre la langue dominante des élèves et la langue d'enseignement	43
Tableau 14. Effets du traitement (échantillon complet avec témoins)	44
Tableau 15. Résultats en mathématiques du groupe de traitement par groupe mathématique	46
Tableau 16 . Effets du traitement selon le genre	48
Tableau 17 . Effets du traitement selon la langue d'enseignement.....	49
Annexe 18 . Moyennes initiales et finales pour le groupe de traitement des élèves sérieux.....	50
Figure 19 . Effets du traitement selon la région	51
Tableau 20. Coûts totaux de l'année scolaire	52
Tableau 21. Coûts de l'Année scolaire par composante du programme	54
Tableau 22. Coûts d'Année Scolaire par catégorie de ressources	54
Tableau 23. Répartition des coûts pour l'Année scolaire par composante du programme	55
Tableau 24. Coûts estimés pour le scénario 1	56
Tableau 25. Coûts estimés pour le scénario 2	57
Tableau 26. Coûts estimés pour le scénario 3	57
Tableau 27. Coûts estimés pour le scénario 4	58
Tableau 28. Inventaire illustratif de l'impact.....	59
Tableau 29. Impacts obtenus pour 100 dollars investis par élève	59

Annexe D1 . Sélection dans les groupes de lecture-écriture <i>Ndaw Wune</i> au début de l'étude	D-5
Tableau D2 . Équivalence de base entre les élèves perdus de vue et les élèves appariés.....	D-6
Tableau D3 . Équivalence de base pour un échantillon apparié	D-7
Tableau D4 . Équivalence de base sans les élèves sérères	D-8
Tableau D5 . Modèle à effets aléatoires	D-9
Tableau D6 . Échantillon complet avec variables de contrôle supplémentaires.....	D-10
Tableau D7 . Échantillon apparié	D-11
Tableau D8 . Échantillon complet sans les élèves sérères	D-13
Tableau E1 . Fiabilité inter-évaluateurs par sous-test de compétences en lecture-écriture pour l'échantillon	E-14

Liste des acronymes

AIR	American Institutes for Research
ANOVA	Analyse de variance
ARED	Associé en recherche et en éducation pour le développement
ASER	Rapport annuel sur l'état de l'éducation
DID	Méthode des doubles différence
EGMA	Évaluation des mathématiques dans les premières années
EGRA	Évaluation des acquis en lecture-écriture au début du cycle primaire
F CFA	Franc de la Communauté financière africaine
PIB	Produit intérieur brut
IEF	Inspection de l'éducation et de la formation
IRB	Comité d'éthique de la recherche
IRR	Fiabilité inter-évaluateurs
ITT	Intention de traiter
LE	Langue d'enseignement
PRFI	Pays à revenu faible et intermédiaire
MEN	Ministère de l'Éducation nationale
MOHEBS	<i>Modèle harmonisé d'enseignement bilingue au Sénégal</i>
RCM	Modèle de coûts par ressources
RCT	Essai contrôlé randomisé
TaRL	Enseignement adapté au niveau de l'élève
TdR	Termes de référence
USD	Dollar américain

Remerciements

Les auteurs tiennent à remercier la Fondation Gates pour le financement de la présente étude. Ils remercient tout particulièrement Clio Dintilhac et Penelope Bender, de la Fondation Gates, pour la qualité de leurs commentaires et observations sur la conception de l'étude, les outils de collecte des données, la méthodologie et le rapport, ainsi que pour les éléments de contextualisation relatifs au Sénégal. Cette évaluation a également bénéficié de l'appui décisif de l'Associé en recherche et en éducation pour le développement (ARED), qui a bien voulu fournir des informations sur le programme *Ndaw Wune* et son cadre opérationnel, ainsi que toutes les données afférentes aux coûts. Nous adressons nos remerciements appuyés à nos consultants, le Professeur Mangary Ka et M. Alioune Badara Diop, ancien Directeur de l'INEADE, pour leur soutien constant et leur disponibilité. Nous remercions par ailleurs Dalberg Research, notre partenaire lors de la collecte de données pour les besoins de la présente évaluation, dont les équipes ont assuré, avec efficacité, la traduction et la programmation des outils de collecte de données, réalisé l'ensemble des évaluations des apprenants, et pris en charge la coordination logistique ainsi que le nettoyage des données, en dépit d'un calendrier de projet particulièrement serré. Nous remercions enfin les nombreux élèves, enseignants et directeurs d'école qui ont prêté leur concours à cette démarche évaluative, en facilitant les évaluations des apprenants. Nous nourrissons l'espoir que les informations recueillies grâce à leurs contributions contribueront à renforcer le programme *Ndaw Wune* et, de façon plus globales, l'éducation au Sénégal.

Résumé

Le Sénégal est confronté à des défis persistants dans le domaine de l'enseignement des compétences fondamentales en lecture-écriture et mathématiques, 13 % seulement des élèves atteignant le niveau minimum requis en lecture et 34 % en mathématiques à la fin du cycle primaire (Suivi des impacts sur les résultats d'apprentissage, n.d.). Bien que le taux d'achèvement du cycle primaire se soit amélioré, passant de 17 % en 2000 à 51 % en 2020, les progrès stagnent depuis 2015 et des disparités persistent, notamment en milieu rural et au niveau des ménages à faibles revenus. Les enseignants ne maîtrisent pas toujours les pratiques pédagogiques requises pour enseigner le programme basé sur les compétences, et beaucoup n'ont pas reçu le guide de l'enseignant en arts du langage ou n'en ont reçu qu'une version numérique (Tous les enfants du Sénégal sont nés pour apprendre, s.d.). Par ailleurs, le système éducatif sénégalais est depuis longtemps confronté au défi de la langue d'enseignement. Même si la quasi-totalité des enfants parlent une langue nationale et non le français avant d'entrer à l'école, cette dernière reste la langue d'enseignement dominante, héritage de la domination coloniale.

Depuis les années 1970, le Sénégal a expérimenté plusieurs approches d'éducation bilingue, et la loi d'orientation sur l'éducation de 2002 a reconnu les langues nationales comme langues potentielles d'enseignement. Plus récemment, le Modèle harmonisé d'enseignement bilingue au Sénégal (MOHEBS) a été introduit dans le but d'institutionnaliser l'éducation bilingue au primaire, avec un horizon de généralisation à l'échelle nationale fixé à 2028. L'enseignement bilingue des arts du langage est actuellement déployé dans 12 des 14 régions administratives du Sénégal, tandis que celui des mathématiques a commencé dans sept régions au cours de l'année scolaire 2025-2026. Malgré les réformes politiques, l'utilisation des langues nationales dans les salles de classe demeure inégale. Les enseignants manquent souvent de formation suffisante en pédagogie bilingue et la disponibilité des supports pédagogiques en langues nationales reste hétérogène (Kébé, s.d. ; Nakamura et al., 2024).

Dans ce contexte, l'Associé en recherche et en éducation pour le développement (ARED) met en œuvre *Ndaw Wune* (« Réussite pour tous »), un programme structuré de rattrapage scolaire ayant pour objet d'apporter un appui pédagogique aux élèves ayant des difficultés d'apprentissage au CP (cours préparatoire) et au CE1 (cours élémentaire première année), en vue de renforcer les compétences de base en lecture-écriture et mathématiques. Le modèle *Ndaw Wune* est fondé sur l'enseignement des langues nationales et sur la mise en place de groupes d'apprentissage différenciés selon le niveau, avec des séquences pédagogiques ciblées élaborées sur la base d'une évaluation des compétences des élèves. Dans ce dispositif, les enseignants dispensent un enseignement direct à un groupe d'élèves d'un même niveau, centré sur une compétence à la fois, tandis que les autres groupes sont engagés dans des activités d'apprentissage autonome. Avec l'appui d'Echidna Giving, d'une fondation familiale anonyme et de la Fondation Gates, et en coordination avec le MEN, ARED

a conçu deux déclinaisons du programme : (a) *Ndaw Wune Année Scolaire*, reposant sur des sessions extrascolaires de rattrapage tenues les lundis, mercredis et vendredis après-midi ; et (b) *Ndaw Wune Vacances*, exécuté durant les grandes vacances (juillet–octobre). Le MEN expérimente actuellement un troisième programme de rattrapage, dénommé « Remédiation harmonisée » encore en cours d’élaboration. Ce programme, qui reprend certains des principes de *Ndaw Wune*, prévoit pour l’ensemble des élèves une heure de rattrapage en mathématiques et une heure en arts du langage comprises dans les quatre heures hebdomadaires consacrées au rattrapage pendant les heures de cours officielles (les mardis et jeudis après-midi).

Pour fournir des données probantes essentielles permettant d’orienter la mise en œuvre et une éventuelle extension de *Ndaw Wune*, l’Institut American de Recherche[®] (AIR[®]) a mené une évaluation indépendante, quasi expérimentale et longitudinale, avec deux cycles de collecte de données, pour mesurer l’impact du programme *Ndaw Wune Année Scolaire* sur les résultats d’apprentissage fondamentaux. Cette évaluation d’impact s’est appuyée sur une méthodologie mixte et une analyse des coûts et du rapport coût-efficacité du programme *Ndaw Wune Année Scolaire*.

Conception de l’évaluation

Analyse DID avec appariement

Pour déterminer l’impact du programme *Ndaw Wune Année Scolaire* au cours de l’année scolaire 2024-2025, nous avons utilisé la méthode des doubles différences (DID) combinée à une méthode d’appariement. Nous avons procédé à l’analyse des impacts de *Ndaw Wune Année Scolaire* pour la cohorte 2024-2025, dont la mise en œuvre pédagogique s’est déroulée de décembre 2024 à mai 2025. Notre démarche méthodologique a consisté, dans un premier temps, à constituer des groupes de comparaison statistiquement similaires grâce à la technique d’appariement par score de propension. Une fois constitués des groupes statistiquement comparables au niveau des écoles, et après prise en compte des différences entre les deux groupes au début du programme, ainsi que des trajectoires d’évolution des compétences des élèves dans le temps au regard d’un scénario « habituel », tel que reflété par le groupe de comparaison, nous estimons les impacts causaux attribuables au programme *Ndaw Wune Année Scolaire*.

Notre approche méthodologique, présentée au tableau 1, procède d’une approche en deux étapes. Dans un premier temps, ARED a procédé à la sélection des écoles bénéficiaires du programme *Ndaw Wune* au titre de l’année scolaire 2024-2025. Dans les quatre régions concernées (Diourbel, Kaolack, Matam et Saint-Louis) abritant les écoles bénéficiaires (Diourbel, Kaolack, Matam et Saint-Louis), nous avons constitué un groupe d’« écoles de comparaison » en fonction de leur proximité géographique et d’un appariement par plus proche voisin. Idéalement, l’appariement aurait également intégré des variables additionnelles observables, notamment la langue d’enseignement. Cependant, ces informations n’étaient pas disponibles avant le démarrage de l’étude. Il en a résulté un déséquilibre linguistique entre les écoles : le groupe de comparaison ne comptait qu’une

école sérère, contre douze dans le groupe de traitement. Par ailleurs, le groupe de comparaison comprenait 23 écoles wolof, contre 12 dans le groupe de traitement, et 11 écoles pulaar dans chacun des deux groupes. L'appariement par plus proche voisin a permis de retenir 35 paires d'écoles (traitées et appariées), soit un échantillon total de 70 écoles.

Dans un deuxième temps, après avoir dressé la liste des écoles retenues pour le traitement et celles sélectionnées pour la comparaison, AIR a constitué un groupe d'élèves de comparaison pour reproduire le processus de sélection des élèves observé dans les écoles de traitement. Dans ces écoles, ARED a administré aux élèves de CP et de CE1 un test de lecture, d'écriture et de mathématiques. Les élèves ayant obtenu un score inférieur à un seuil prédéterminé (par exemple, la capacité de lire au moins 10 mots simples et familiers) ont été considérés comme éligibles au programme. Seize élèves ont ensuite été sélectionnés aléatoirement par école traitée pour constituer l'échantillon d'élèves traités. Pour nous assurer d'une comparabilité étroite entre les élèves de comparaison de notre échantillon et ceux sélectionnés pour bénéficier d'un rattrapage dans les écoles *de Ndaw Wune*, nous avons suivi le processus d'ARED en administrant le même test aux élèves des écoles de comparaison. Après identification des 35 élèves éligibles dans chaque école de comparaison, nous avons procédé à un tirage aléatoire de 16 élèves pour constituer l'échantillon de comparaison.

Une fois les élèves de comparaison et ceux du groupe de traitement identifiés, nous avons recueilli des données de référence au début de l'année scolaire (novembre-décembre 2024) et des données de suivi à la fin de l'année scolaire (mai-juin 2025). À la fin de l'étude, 23 % de l'échantillon initial n'a pu être retrouvé. Toutefois, l'attrition différentielle entre le groupe de traitement et le groupe de comparaison est restée faible, à 5 %. Des élèves de remplacement ont été sélectionnés à la fin de l'étude (sur la base de critères similaires à ceux mentionnés ci-dessus) pour garantir une taille d'échantillon suffisamment large à la fin de l'étude pour détecter des effets significatifs. Nous présentons des tests de robustesse uniquement pour les élèves qui étaient présents au début et à la fin de l'étude, ainsi que pour l'échantillon complet, y compris les élèves de remplacement.

1 . Approche méthodologique

Résumé du dispositif de recherche de Ndaw Wune

AIR a mené l'évaluation de Ndaw Wune au cours de l'année scolaire 2024–2025, en utilisant une évaluation d'impact quasi expérimentale, une analyse qualitative des données de suivi, ainsi qu'une analyse rigoureuse de la rentabilité/coût efficacité.

Méthodologie

Évaluation d'impact

L'évaluation d'impact a utilisé une méthode d'appariement (matching) et une approche de différences en différences afin de comparer les compétences en lecture et en mathématiques des élèves des groupes de traitement et de comparaison, de novembre 2024 à juin 2025.

 **Élèves**
1,109 en situation de référence
1,059 en situation finale

 **Écoles**
35 Traitement
35 Comparaison

 **Commentaires des enseignants**

Analyse des coûts

AIR a intégré une méthodologie de calcul des coûts basée sur les intrants et les activités, du point de vue du programme, pour chacune des trois versions de Ndaw Wune.

 **Personnel**  **Déplacements**  **Frais généraux**  **Matériels et équipements**

Analyse des données de suivi

Les chercheurs d'AIR ont utilisé les commentaires fournis par ARED afin d'apporter des éclairages sur la mise en œuvre du programme, en complément de l'évaluation d'impact.

Coût efficacité

AIR a mené une analyse coût conséquence afin de comparer les résultats de l'analyse des coûts avec ceux de l'évaluation d'impact, dans le but de déterminer le coût relatif de Ndaw Wune au regard des impacts générés.

Analyse des coûts

AIR a procédé à une analyse rétrospective des coûts du programme de rattrapage scolaire *Ndaw Wune*. Autrement dit, nous avons collecté les coûts du point de vue du programme, y compris tous les coûts directement engagés par l'entité chargée de la mise en œuvre. Notre équipe a produit des estimations de coûts pour l'itération 2024-2025 du programme : *Ndaw Wune Année Scolaire*, sur la base des coûts observés au regard des régions, des effectifs d'élèves et des écoles desservies par le programme pendant cette période.

AIR a adopté une approche hybride d'analyse des coûts, articulant une méthode par intrants et une méthode orientée activités. L'approche par intrants, reconnue comme une bonne pratique mais fortement consommatrice de ressources (Danks & Kendziora, 2021 ; Levin et al., 2018), repose sur un processus systématique d'identification et de quantification de l'ensemble des intrants du programme (p. ex. le nombre d'heures-personnes par rôle et par activité, le type et le nombre de matériaux mobilisés, etc.), puis sur l'affectation d'un coût unitaire à chaque intrant pour en estimer le coût total. Pour optimiser le rapport coût-efficacité et de réduire la charge de travail liée à la collecte de données pour le personnel et les participants, la collecte de données pour l'étude des coûts a été circonscrite aux documents programmatiques et financiers transmis par ARED, ainsi qu'à des entretiens avec les principaux membres du personnel d'ARED. Sur la base des informations disponibles, l'approche par intrants a été appliquée aux ressources pour lesquelles des données détaillées étaient accessibles (quantités et coûts unitaires), notamment le nombre de cahiers d'exercices destinés aux élèves et les frais de déplacement remboursés aux enseignants participant aux formations.

Pour les intrants du programme pour lesquels nous ne disposions pas de suffisamment d'informations désagrégées, au niveau des « intrants », nous avons complété l'analyse par une approche fondée sur les activités. Cette approche consiste à recenser, pour chaque activité du programme, les intrants nécessaires, puis à en estimer la valeur à partir des budgets du programme, des pièces et justificatifs de dépenses, des registres d'activités et d'entretiens avec des informateurs clés (Brander et al., 2024). En d'autres termes, au lieu de recueillir des données sur le nombre d'heures consacrées par chaque membre du personnel d'ARED à chaque activité dans le cadre du programme *Ndaw Wune*, puis de multiplier ce nombre par le coût horaire de la main-d'œuvre pour déterminer les coûts de personnel pour chaque composante du programme, l'approche par activités s'appuie sur les descriptions des activités du programme, les données sur les salaires annuels et les entretiens avec le personnel d'ARED pour ventiler, au prorata, les coûts de personnel estimés entre les activités du programme. Cette approche hybride est efficace pour produire des données rigoureuses et exploitables sur les coûts, pouvant informer la prise de décision relative à la mise à l'échelle du programme.

Analyse qualitative des données de suivi

Pour compléter l'évaluation, ARED a communiqué des données qualitatives issues de son système de suivi, qui couvre les 35 écoles participant au programme. Ces données comprenaient de brefs informations auprès de 54 enseignants sur la période de décembre 2024 à avril 2025. Les enseignants de *Ndaw Wune* sont des agents du secteur public qui perçoivent une allocation supplémentaire pour mettre en œuvre le programme. Le Sénégal compte de nombreuses catégories d'enseignants ; les enseignants de *Ndaw Wune* comprennent des fonctionnaires, des enseignants contractuels et des enseignants communautaires, présentant des niveaux distincts de formation, d'expérience et de rémunération. Les chercheurs d'AIR ont exploité les commentaires des enseignants pour en extraire des informations utiles sur le programme, qui pourraient aider ARED et le MEN dans la réflexion relative à une éventuelle mise à l'échelle du programme. Notre équipe a suivi un processus systématique pour d'analyse qualitative des commentaires des enseignants. À l'aide du logiciel de recherche qualitative NVivo, l'ensemble des commentaires a d'abord été importé puis structuré selon (a) le superviseur, (b) la région et (c) l'école. Les commentaires ont ensuite fait l'objet d'un codage inductif en fonction des thèmes émergents tels que « absences des élèves » ou « manque de matériel ». Une fois le processus de codage thématique terminé, nous avons utilisé les résultats pour fournir des informations utiles à la mise en œuvre du programme. Il convient toutefois de souligner que, les commentaires des enseignants étant autodéclarés et de nature anecdotique, ils servent uniquement à identifier des thèmes pertinents à considérer dans une perspective d'extension, et non nécessairement à expliquer les résultats de l'évaluation d'impact.

Résultats

Résultats de l'évaluation d'impact

Notre évaluation est adossée aux principes théoriques de la « Simple View of Reading » (Gough & Hoover, 1989), plus récemment conceptualisée sous l'intitulé « Cognitive Foundations of Reading Acquisition » (Hoover & Tunmer, 2020). Dans ce cadre, la

compréhension de base en lecture dépend de l'intégration des compétences linguistiques, des capacités de traitement et de la reconnaissance des mots (ou décodage).

Dans l'ensemble, l'analyse montre que *Ndaw Wune* a eu un impact positif sur les compétences fondamentales en lecture-écriture. Les élèves du groupe traité ont obtenu des résultats significatifs et positifs dans plusieurs catégories de lecture/écriture, notamment le décodage (0,42 *écart-type*)¹, la fluidité de lecture à voix haute (0,32 *écart-type*)² et la compréhension en lecture (0,14 *écart-type*)³. *Ndaw Wune* a également eu un impact plus faible, mais toujours positif, sur les performances en mathématiques. Deux des six sous-tests de mathématiques affichent des effets positifs et significatifs : le sens des nombres (0,19 *écart-type*)⁴ et la soustraction (0,20 *écart-type*)⁵. Tous les autres sous-tests ne mettent pas en évidence de résultats significatifs.

Nous avons également analysé l'hétérogénéité des effets selon le genre, la langue d'enseignement et la région. Il convient toutefois de préciser que ces analyses n'étaient pas suffisamment puissantes pour mettre en évidence des effets statistiquement significatifs et que les résultats doivent donc être considérés comme suggestifs. Les effets apparaissent globalement plus marqués chez les garçons. Bien que les performances des garçons et des filles soient positivement influencées par *Ndaw Wune*, les garçons présentent des effets positifs et significatifs sur trois sous-tests de lecture-écriture, contre deux pour les filles, la différence portant sur la compréhension écrite. Les résultats en mathématiques sont plus mitigés : les garçons montrent des effets positifs en addition et soustraction, tandis que les filles présentent un effet positif sur le sens des nombres et d'un effet négatif sur le sous-test de comptage.⁶

Les progrès les plus substantiels sont observés chez les élèves parlant le wolof. Les données indiquent que les wolophones enregistrent généralement de meilleures performances que les pulaarophones sous-tests de lecture-écriture que les pulaarophones.⁷ Bien que les élèves wolophones présentent un effet légèrement négatif au niveau de la compréhension orale (p

¹ Le décodage est une construction composée de trois mesures de décodage. Les connaissances des élèves en traitement sur les sons des lettres sont passées de 70,75 sur 100 au début de l'étude à 83,52 à la fin, le décodage des mots inventés est passé de 5,99 sur 50 au début à 11,85 à la fin, et le décodage des mots familiers est passé de 6,74 sur 50 au début à 12,87 à la fin.

² La fluidité de la lecture à voix haute des élèves ayant suivi le traitement est passée de 5,57 sur 50 au début de l'étude à 13,12 à la fin.

³ La compréhension écrite des élèves ayant suivi le traitement est passée de 0,68 sur 6 à 1,35 à la fin du programme.

⁴ Le sens des nombres des élèves ayant suivi le traitement est passé de 4,48 sur 12 au début de l'étude à 5,84 à la fin.

⁵ La capacité de soustraction des élèves du groupe expérimental est passée de 5,12 sur 12 au début de l'étude à 7,26 à la fin.

⁶ Au début de l'étude, les filles du groupe traité avaient des scores statistiquement significativement plus élevés que les garçons du groupe traité pour deux des trois résultats de décodage, l'addition et la soustraction. L'ampleur plus importante de l'effet du traitement chez les garçons pourrait s'expliquer par le fait que ceux-ci ont plus de marge de progression que les filles par rapport à leurs valeurs de référence. Des études connexes ont mis en évidence des effets mixtes selon le contexte. Par exemple, dans l'étude sur le programme de rattrapage en lecture et écriture en Afrique du Sud, les auteurs ont noté que les filles surpassaient les garçons au début et à la fin de l'étude et que l'avantage des filles se maintenait même dans leurs modèles de régression multivariée (Fleisch et al., 2017). L'évaluation d'un programme de rattrapage au Niger n'a pas révélé de différences entre les sexes dans les effets du traitement (Maruyama & Kurosaki, 2021).

⁷ Nous avons examiné la correspondance entre la langue d'enseignement et la ou les langues connues par les élèves comme explication potentielle des meilleurs résultats des élèves wolofs aux évaluations. Cependant, dans toutes les langues, nous constatons une forte correspondance (de l'ordre de 96 à 100 %) entre la langue d'enseignement et les langues dominantes des enfants. De plus, en termes de niveaux de référence, les élèves wolophones ont obtenu des scores plus élevés dans la plupart des sous-compétences que les locuteurs pulaar, ce qui semble s'être poursuivi jusqu'à la fin de l'étude.

$< 0,10$)⁸, ils affichent des gains positifs plus substantiels en décodage (0,49 *écart-type* contre 0,33 *écart-type* pour les locuteurs pulaar) et un effet de traitement de 0,53 *écart-type* sur la fluidité de lecture à voix haute, tandis que l'effet de traitement pour les locuteurs pulaar en matière de fluidité de lecture à voix haute n'est pas statistiquement significatif.

Au niveau régional, les résultats les plus marqués sont observés à Kaolack. Dans cette région, trois des cinq sous-tests de lecture-écriture (conscience phonologique, décodage et fluidité de lecture à voix haute) présentent des résultats significativement positifs, contre un seul à Saint-Louis et aucun à Matam. Comme indiqué, l'analyse n'était pas suffisamment puissante pour détecter les effets au niveau des sous-groupes. Ainsi, l'absence d'impact à Matam pourrait s'expliquer par une puissance statistique insuffisante. Aucune des régions prise individuellement ne montre d'effet positif et significatif du traitement sur la composante mathématiques.

Coûts-efficacité

Nous estimons le coût total du programme *Ndaw Wune Année Scolaire* à 251 141 271 F CFA (433 002 USD), soit 83 714 F CFA (144 USD) par élève pour l'année scolaire 2024-2025. Nous constatons que la majeure partie des coûts est consacrée aux activités d'enseignement direct (62 %), puis à la gestion, à la planification et à l'évaluation (23 %). La moitié environ des coûts du programme (51 %) est consacrée aux frais de personnel directs, tandis que près d'un cinquième (20 %) des coûts est réservé à la prise en charge des déplacements liés aux formations, aux journées pédagogiques et aux évaluations. D'après l'analyse coût-efficacité, un investissement de **100 USD par élève** se traduit par une augmentation modérée de l'effet sur la connaissance des sons des lettres, le décodage des mots et la fluidité de lecture à voix haute, ainsi que par une légère augmentation de l'effet sur la compréhension en lecture et les résultats en mathématiques.

Nous avons également établi des projections de coûts pour quatre scénarios d'extension du programme *Ndaw Wune Année Scolaire* (extension à 50 000 élèves dans les six IEF) :

- **Scénario 1** : Aucune modification n'est apportée au calendrier, la formation initiale des enseignants ou les journées pédagogiques. Les directeurs d'école assument le rôle de superviseurs. Le coût estimé pour le scénario 1 est de 767 650 752 F CFA (1 323 536 USD) et le coût par élève est de 15 353 F CFA (26 USD).
- **Scénario 2** : Le calendrier quantitatif demeure inchangé. La formation initiale des enseignants est réduite à quatre demi-journées. Les indemnités versées aux superviseurs et aux inspecteurs sont supprimées. Le coût estimé pour le scénario 2 est de 638 388 975 F CFA (1 100 671 USD) et le coût par élève est de 12 768 F CFA (22 USD).
- **Scénario 3** : L'enseignement est entièrement dispensé pendant les heures scolaires officielles. La formation initiale des enseignants est réduite à quatre demi-journées et les indemnités versées aux superviseurs, aux inspecteurs et aux enseignants sont

⁸ Au départ, les élèves des écoles wolofs ont obtenu des scores élevés au test de compréhension orale (4,4 sur 6 contre 3,4 sur 6 pour les élèves du groupe témoin). Bien que les deux groupes aient progressé entre le début et la fin de l'étude, un effet plafond a peut-être empêché les élèves du groupe expérimental de progresser à un rythme identique ou supérieur à celui des élèves du groupe témoin, ce qui a conduit à une estimation négative de l'impact.

supprimées. Le coût estimé du scénario 3 est de 360 055 641 F CFA (620 786 USD) et le coût par élève est de 7 201 F CFA (12 USD).

- **Scénario 4** : Toutes les hypothèses du scénario 3 sont maintenues, et les élèves utilisent le matériel existant, ce qui supprime les coûts liés à l'impression et à la livraison des supports pédagogiques. Le coût estimé pour le scénario 4 est de 263 343 808 F CFA (454 041 USD) et le coût par élève est de 5 267 F CFA (9 USD).

Analyse qualitative des résultats des données de suivi

Le programme *Ndaw Wune* a été conçu pour offrir deux heures de rattrapage, trois jours par semaine. Chaque session prévoit environ 60 minutes consacrées à la lecture et à l'écriture (y compris la chanson de l'alphabet), 30 minutes aux mathématiques, 15 minutes aux jeux éducatifs de mathématiques et de lecture-écriture, et environ 15 minutes à la chanson de l'alphabet, à la récréation et à l'attribution des devoirs à la fin du cours. Le modèle pédagogique divise les élèves en trois groupes pour la lecture-écriture : un pour les élèves pouvant identifier moins de 10 lettres, un pour les élèves qui peuvent identifier au moins 10 lettres mais moins de 10 syllabes, et un autre pour les élèves pouvant lire (décoder) au moins 10 syllabes mais moins de 10 mots. Il prévoit deux groupes pour l'apprentissage des mathématiques : un pour ceux qui peuvent effectuer correctement moins de cinq comparaisons de nombres et un pour les élèves pouvant effectuer correctement les cinq comparaisons de nombres.

L'analyse qualitative des données de suivi d'ARED indique que le programme a apporté un soutien substantiel aux élèves, et ce de plusieurs manières. Les enseignants ont rapporté que la mise à disposition de guides pédagogiques, de cahiers d'exercices et de cartes alphabétiques a contribué à améliorer la qualité de l'enseignement et à accroître l'efficacité des activités d'apprentissage. La structure du programme, fondée sur la remédiation et des passages flexibles entre groupes, aurait permis aux élèves de bénéficier d'un appui ciblé en fonction de leurs progrès et de leurs besoins. Les enseignants ont également signalé certaines difficultés de mise en œuvre. Il s'agit d'informations anecdotiques fournies par les enseignants eux-mêmes et mentionnées par moins de 1 à 7 % des enseignants dans chaque cas. Les principales difficultés récurrentes qui ressortent de ces données sont les suivantes : un effectif d'élèves perçu comme élevé rendant la gestion de la classe plus difficile ; des taux d'absences élevés ; l'indisponibilité ou l'absence de certains matériels nécessaires à l'enseignement ; des inexactitudes dans les supports ; et des supports de mathématiques jugés difficiles à utiliser.

Limites

La présente étude présente certaines limites.

1. Les données disponibles au niveau des écoles pour procéder à l'appariement entre écoles de traitement et écoles témoins étaient limitées. Idéalement, l'appariement aurait pu être fait sur la base de caractéristiques observables, notamment les taux d'assiduité des élèves et des enseignants, les taux d'abandon scolaire, l'existence de cantines scolaires et d'autres ressources éducatives et, surtout, la langue d'enseignement. Cependant, la plupart de ces données n'étaient pas disponibles avant la collecte des données lors de l'appariement. Ainsi, bien que l'échantillon apparié au niveau des écoles repose sur le nombre limité de variables disponibles, les appariements pourraient différer selon de

nombreuses autres caractéristiques, observables comme non observables. Par ailleurs, faute d'accès aux informations relatives à la langue d'enseignement des écoles avant la réalisation de l'appariement, l'échantillon final comprenait seulement 12 écoles de traitement en sérère et une seule école de comparaison en sérère (les données sur la langue d'enseignement n'étant disponibles pour les écoles de comparaison qu'au moment où les enquêteurs se sont rendus dans les établissements pour la collecte). Cette taille d'échantillon n'est pas suffisante pour produire des comparaisons fiables entre les groupes de traitement et de comparaison pour la langue sérère ; en conséquence, la présente étude ne permet pas de tirer des conclusions pour le sous-échantillon sérère.

2. Nous disposons d'une puissance statistique limitée pour détecter des effets hétérogènes. Bien que la taille globale de l'échantillon soit importante, notre dispositif d'évaluation repose sur un nombre restreint d'établissements scolaires, ce qui limite notre capacité à identifier, avec un degré de précision suffisant, d'éventuelles hétérogénéités dans les estimations d'impact.
3. Une hypothèse centrale de la méthodologie des doubles différences (DID) est que les élèves du groupe de traitement et ceux du groupe de comparaison suivaient des parcours identiques avant le démarrage de l'intervention (autrement dit, en l'absence de *Ndaw Wune*, les élèves du groupe de traitement et ceux du groupe de comparaison auraient présenté des tendances similaires dans leurs résultats). Cette hypothèse est généralement examinée à partir de données collectées à plusieurs moments avant le début de la mise en œuvre. Il est fréquent (comme dans le cas présent) que la collecte de données antérieures ne soit pas possible. Dans ce cas, la stratégie d'appariement ainsi que les tests d'équivalence à la ligne de base contribuent à atténuer cette limitation.
4. Dans le cadre de l'analyse des coûts, bien que notre approche hybride combinant la méthode par intrants et la méthode par activités, soit à la fois pertinente et efficace pour les besoins de la présente étude, elle implique un arbitrage en termes de précision. Pour les activités du programme pour lesquelles nous ne disposons pas de données détaillées sur les dépenses, nous avons estimé les coûts sur la base des descriptions des activités, des termes de référence (TdR) et des entretiens réalisés avec le personnel organisationnel d'ARED. Nos données sur les coûts pourraient également être sujettes à un biais de rappel de la part des responsables de la mise en œuvre du programme. La triangulation entre les données documentées sur les coûts et les données issues des entretiens a permis d'atténuer les inconvénients associés à ce biais de rappel.

Conclusions et recommandations

L'évaluation a montré que le programme *Ndaw Wune Année Scolaire* entraîne des améliorations significatives des compétences fondamentales en lecture-écriture (décodage : 0,42 *écart-type*, fluidité de lecture à voix haute : 0,32 *écart-type*, compréhension écrite : 0,14 *écart-type*) et des gains modestes en mathématiques (sens des nombres : 0,19 *écart-type*, soustraction : 0,20 *écart-type*) chez les élèves en difficulté des premières années du primaire. Ces résultats constituent des éléments probants solides indiquant qu'un enseignement de rattrapage ciblé, axé sur des compétences spécifiques, dispensé dans des langues maîtrisées par les enfants et organisé en groupes de niveau, peut contribuer à résorber les difficultés fondamentales d'apprentissage au Sénégal. Les décideurs publics devraient, en conséquence,

faire de la mise à l'échelle de tels modèles de remédiation une priorité, en particulier dans les régions et auprès des populations présentant les résultats scolaires sont les plus faibles.

Nous formulons ci-après des recommandations clés pour les prochaines étapes, dans la perspective d'une mise à l'échelle du programme. Ces recommandations peuvent être mises en œuvre dans le cadre du programme *Ndaw Wune*, à travers les dispositifs de programmation et d'appui du MOHEBS, par d'autres programmes d'exécution relevant du MOHEBS, ou encore par d'autres entités externes.

Pour le perfectionnement professionnel des enseignants et d'enseignement en classe :

- 1. Élargir l'enseignement explicite de la phonétique aux classes ordinaires.** Le programme *Ndaw Wune* a démontré que le renforcement des compétences de mise en correspondance graphophonologique par un enseignement explicite de la phonétique, combiné à des opportunités structurées visant à améliorer l'automatisme (rapidité et précision dans l'établissement des correspondances graphème-phonème aux niveaux sous-lexical et lexical), s'avère particulièrement efficace pour améliorer les sous-compétences de décodage et la compréhension en lecture. Au regard des résultats nettement positifs observés pour ces sous-compétences dans le cadre de l'évaluation de *Ndaw Wune*, nous recommandons également que le **MOHEBS** s'appuie sur ces enseignements et envisage l'intégration systématique d'un enseignement explicite de la phonétique dans les langues sénégalaises dans les classes sur l'ensemble du territoire, pratique qui demeure absente dans la majorité des salles de classe (plus de 73 % selon Nakamura et al., 2024) au Sénégal.
- 2. Veiller à l'expansion et à la mise à l'échelle des composantes de pédagogie structurée pour l'enseignement des compétences fondamentales dans les salles de classe à travers le Sénégal.** La combinaison d'une planification fondée sur une progression structurée des apprentissages (*scope and séquence*), de guides pédagogiques clairs et scénarisés (dans une certaine mesure), de supports d'enseignement et d'apprentissage, en particulier de cahiers d'exercices pour chaque enfant, a démontré, au regard des résultats de la présente évaluation, un impact significatif sur le développement des compétences ciblées et devrait être garantie dans les programmes de formation des enseignants et suivie par le biais d'un contrôle de la fidélité d'exécution dans les programmes d'apprentissage fondamental à travers le pays. Compte tenu du manque de cahiers d'exercices et de matériel d'enseignement et d'apprentissage dans les salles de classe, de l'exposition limitée à l'écrit en dehors du cadre scolaire, ainsi que de l'état souvent dégradé des tableaux noirs — difficiles à lire dans des salles de classe de grande taille — la mise à disposition d'un cahier individuel pour chaque élève constitue une composante essentielle des dispositifs de pédagogie structurée nécessaires à l'amélioration des apprentissages fondamentaux.
- 3. Envisager un renforcement supplémentaire de l'appui pédagogique en compréhension orale dans les langues nationales.** Les résultats de l'évaluation n'ont mis en évidence aucune amélioration significative des compétences langagières orales, pourtant constitutives d'un des piliers essentiels du développement des compétences fondamentales en lecture-écriture. Par ailleurs, ce résultat ne saurait être attribuée à un effet plafond au niveau de la situation initiale. Par conséquent, *Ndaw Wune* et les autres programmes nationaux doivent renforcer la compréhension

orale et les compétences lexicales orales, même s'ils enseignent déjà dans des langues que les enfants connaissent et comprennent. En d'autres termes, même lorsque les enfants reçoivent un enseignement dans leur propre langue, il est nécessaire de consacrer un temps d'enseignement considérable au renforcement du vocabulaire et des connaissances linguistiques dans l'ensemble du programme, et de ne pas se concentrer uniquement sur l'enseignement de la phonétique et du décodage.

4. **Investir dans l'amélioration des composantes du programme liées aux mathématiques.** L'évaluation d'impact a mis en évidence des difficultés dans l'enseignement des mathématiques. Nous proposons donc quelques approches pour améliorer l'enseignement des bases des mathématiques. Premièrement, l'enseignement devrait suivre une progression explicite au sein de chaque opération, en passant d'une perception des nombres à plusieurs chiffres comme une simple juxtaposition de chiffres isolés vers une compréhension structurée de la valeur positionnelle (par exemple, comprendre que 10 unités équivalent à une dizaine) (Rogers, 2023). Deuxièmement, les enseignants pourraient introduire explicitement des stratégies de calcul efficaces allant au-delà du simple comptage par dénombrement, telles que le passage par la dizaine, les sauts numériques, ou encore les méthodes d'arrondi et d'ajustement, dont l'efficacité pour améliorer la fluidité du calcul mental a été démontrée lorsqu'elles sont modélisées et pratiquées de manière systématique (Kumm & Graven, 2024). Troisièmement, l'intégration de discussions structurées autour des nombres permettrait d'encourager les élèves à expliciter et comparer différentes stratégies de résolution, favorisant ainsi le développement du sens du nombre, de la flexibilité cognitive et des compétences métacognitives grâce à un discours collaboratif. Enfin, le recours intensif à des supports pédagogiques de groupement, tel que des paquets de bâtonnets ou des blocs de base dix fabriqués localement, peut rendre tangibles des concepts abstraits liés à la valeur positionnelle (notamment l'idée que 10 unités constituent une dizaine) et faciliter le regroupement dans les opérations (Resnik, 1987).
5. **Nécessité de renforcer le transfert inter-linguistique et les compétences linguistiques orales en français au regard des objectifs d'éducation bilingue du Sénégal.** L'objectif de *Ndaw Wune* est d'améliorer les compétences en lecture-écriture dans les langues nationales. L'objectif des apprentissages fondamentaux au Sénégal est l'acquisition, par les élèves, de compétences en lecture-écriture dans les langues nationales et le français (Nakamura et al., 2026). Les sciences de la lecture en contexte bilingue montrent clairement que l'apprentissage de la lecture dans une nouvelle langue est fortement tributaire de la capacité des élèves à lire dans leur langue maternelle et de leur niveau de maîtrise orale (ou signée) de la langue additionnelle à apprendre. Un programme de littératie fondamentale efficace au Sénégal devrait donc concevoir des curricula et des supports pédagogiques en tenant compte, dès le début du primaire, des compétences transférables et des fondements cognitifs de l'acquisition de la bilittératie pour faciliter la transition vers le primaire supérieur.

Pour des recherches plus approfondies :

6. **Comprendre les interactions entre le développement du langage et le développement des compétences en mathématiques.** Les recherches futures ainsi que la conception des programmes au Sénégal devraient examiner de façon systématique la manière dont les compétences linguistiques et le choix de la langue d'enseignement influencent la progression des apprentissages en mathématiques. Les recherches devraient analyser dans quelle mesure l'enseignement bilingue ou multilingue, au sein duquel les élèves apprennent les mathématiques à la fois dans leur langue maternelle et en français, influe sur le raisonnement mathématique, le transfert de compétences et les résultats à long terme. Il conviendrait notamment d'évaluer si une maîtrise précoce des mathématiques en wolof ou en pulaar facilite ultérieurement l'apprentissage des mathématiques en français, ou si elle engendre des difficultés lors de la transition vers le vocabulaire et les concepts mathématiques. Les recherches futures devraient également étudier comment les inadéquations linguistiques affectent les résultats en mathématiques, facilite ultérieurement l'apprentissage des mathématiques en français, ou si elle engendre, au contraire, des difficultés lors de la transition vers le vocabulaire et les concepts mathématiques.
7. **Veiller à la collecte des données sur la correspondance entre la langue de l'enseignants.** Avec déploiement d'initiatives d'envergure en langues nationales, le MEN envisage de renforcer la collecte des informations sur la langue des enseignants au sein de MIRADOR (système d'information pour la gestion des enseignants), en appui à l'opérationnalisation de l'éducation bilingue. Nous recommandons de compléter ces données par des données linguistiques relatives aux élèves, de manière à mesurer l'ampleur des correspondances et des inadéquations entre les langues de l'enseignant et celles des élèves.
8. **Réaliser une évaluation du processus pour approfondir les composantes du programme les plus susceptibles d'améliorer les acquis des élèves et d'expliquer les raisons des variations régionales.** Au moment où ARED cherche à affiner et à mettre à l'échelle le programme *Ndaw Wune*, une évaluation correspondante du processus serait extrêmement utile pour identifier les composantes clés du programme qui sont bénéfiques pour les élèves et les enseignants et pour préciser comment le programme peut reproduire les succès et résoudre les problèmes afin de l'optimiser en vue de son extension. Une telle évaluation pourrait analyser la fidélité de la mise en œuvre en comparant les activités prévues, telles que la prestation structurée des cours et l'utilisation des supports pédagogiques, avec ce qui a effectivement été réalisé dans les écoles. Elle pourrait également évaluer la qualité de la prestation au moyen d'observations en classe permettant de déterminer dans quelle mesure les enseignants appliquent les méthodes pédagogiques promues par le programme, la cohérence du coaching et la nature des adaptations opérées dans la pratique. Le fait de recueillir les points de vue des enseignants par le biais d'entretiens, d'enquêtes et de groupes de discussion pourrait aider à mettre en lumière la façon dont les

enseignants perçoivent l'utilité de la formation et des supports pédagogiques, les défis auxquels ils sont confrontés et les conditions qui favorisent ou entravent le plus une mise en œuvre efficace. L'évaluation pourrait également examiner de près l'engagement des élèves et la dynamique de classe pour identifier les pratiques pédagogiques qui conduisent le plus systématiquement à une participation active et à la compréhension. En outre, le suivi de la dose du programme (c'est-à-dire le temps que les élèves passent réellement à suivre les cours *Ndaw Wune*) pourrait contribuer à expliquer les variations d'effets observées selon les écoles et les régions.

9. Assurer un suivi systématique des données relatives à l'assiduité des élèves. Étant donné que les données relatives à l'assiduité que nous avons reçues étaient incohérentes et présentaient plusieurs lacunes, les travaux futurs devraient envisager la mise en place d'un dispositif de suivi systématique et rentable permettant de suivre l'assiduité de chaque élève. Un tel dispositif permettrait d'identifier les causes des absences et de détecter l'absentéisme chronique. Il pourrait s'agir d'un simple outil basé sur des messages texte (SMS) qui alimenterait un dispositif de prise de décision fondé sur des données probantes et ne nécessiterait pas l'accès à Internet ou à un téléphone intelligent. Une telle approche a été testée avec succès dans plusieurs pays, notamment en Afrique de l'Ouest (Mugo & Piper, 2025). À titre d'exemple, un projet pilote au Kenya a montré que le coût de maintenance d'un système centralisé de suivi de l'assiduité par SMS pour 10 000 élèves au niveau central, complété par l'envoi hebdomadaire de SMS personnalisés aux enseignants concernant les élèves en difficulté, coûterait moins de 0,08 USD par enfant et par an (Mugo & Piper, 2025). L'amélioration des systèmes de données permettra également de suivre l'assiduité, l'attrition et les impacts hétérogènes, ce qui permettra une gestion plus réactive des programmes.

10. Mobiliser l'intelligence artificielle (IA) et les technologies éducatives (le cas échéant) pour appuyer la production de supports d'enseignement et d'apprentissage (TLM) et les dispositifs de formation. En matière de formation, de courtes vidéos générées par l'IA pourraient être utilisées pour des sessions fréquentes à faible intensité et pour un accompagnement de soutien. Elles peuvent également constituer un apport complémentaire sur des thématiques difficiles ou des séances pratiques liées aux mathématiques. Il peut également être utile d'utiliser l'IA pour produire des supports additionnels « de secours », conformes à tous les principes scientifiques de la lecture bilingue évoqués ci-dessus. À titre d'exemple, il peut s'agir de :

- a. veiller à ce que chaque élève ait accès au manuel scolaire, que chaque enseignant dispose d'un guide et qu'ils soient alignés. Dans les rares cas où les élèves doivent partager des manuels, la formation des enseignants devrait intégrer des consignes pratiques sur l'animation d'activités de partage.
- b. recourir à l'IA pour créer des images utilisables comme cartes d'histoires, cartes de lettres et autres supports, de manière plus efficiente sur le plan des

coûts, sans dépendre systématiquement des contraintes linguistiques actuelles de l'IA.

- c. veiller à ce que le contenu du manuel scolaire couvre toutes les sous-compétences en lecture-écriture et de langage selon une trajectoire d'apprentissage appropriée—en particulier pour l'acquisition de la bilinguïté et la transition ultérieure vers le français—afin de réduire la fragmentation des supports à imprimer et à harmoniser.

11. Pour les politiques :

12. Importance de commencer par enseigner aux enfants les langues nationales du Sénégal dans le cadre du MOHEBS. Le programme *Ndaw Wune* montre clairement que lorsque l'enseignement est dispensé dans des langues que les enfants utilisent et comprennent, en mettant l'accent sur l'enseignement explicite des sons des lettres et des règles phonétiques, on constate des améliorations significatives dans les capacités de décodage et de compréhension en lecture, même chez les lecteurs les plus en difficulté. Il est important de noter que *Ndaw Wune* est un exemple exceptionnel qui montre que lorsque l'enseignement est dispensé dans des langues que les élèves comprennent, les ressources peuvent être concentrées sur la composante « reconnaissance des mots » de l'apprentissage fondamental de la lecture-écriture. Cette approche, par opposition à l'enseignement directement en français, rend l'acquisition de la lecture-écriture plus efficiente, un temps de classe précieux n'a pas à être mobilisé pour l'enseignement de compétences orales de base et peut être consacré prioritairement aux sons des lettres, à l'assemblage, à la fluidité et au développement de compétences orales de niveau supérieur. En plus du fait que les élèves comprennent la langue d'enseignement, il est également indispensable que la langue de l'enseignant corresponde à celle des élèves. Les données de la présente étude pas mis en évidence de différences significatives entre la langue des élèves et la langue d'enseignement. Toutefois, nous n'avons pas collecté de données sur les compétences linguistiques des enseignants. Des inadéquations importantes persistent encore à l'échelle nationale entre la langue des élèves, celle des enseignants et la langue d'enseignement (Nakamura et al., 2024 ; Diallo & Mbaye, 2025) et doivent être corrigées afin d'améliorer les apprentissages fondamentaux, tant en langues nationales qu'en français.

1. Introduction

Pour améliorer l'apprentissage fondamental au Sénégal, l'Association pour la recherche et l'éducation pour le développement (ARED) appuie l'éducation bilingue dans tout le pays. ARED est actuellement le principal partenaire du gouvernement dans le cadre de l'élaboration et de la mise en œuvre du *Modèle harmonisé d'enseignement bilingue au Sénégal* (MOHEBS), en fournissant une assistance technique liée à l'enseignement dans les langues nationales et à la transition vers le français, notamment en élaborant des manuels scolaires, des guides pédagogiques et des livres d'entraînement à la lecture, et en soutenant le développement professionnel des enseignants. ARED appuie également le développement de *la Remédiation harmonisée*, dirigée par le MEN, qui propose des cours de rattrapage en mathématiques et lecture-écriture dans une centaine de classes de CI et CP, intégrés à la journée scolaire officielle. La connaissance approfondie d'ARED du contexte éducatif et linguistique du Sénégal continue d'ouvrir la voie à des réformes éducatives à l'échelle du pays.

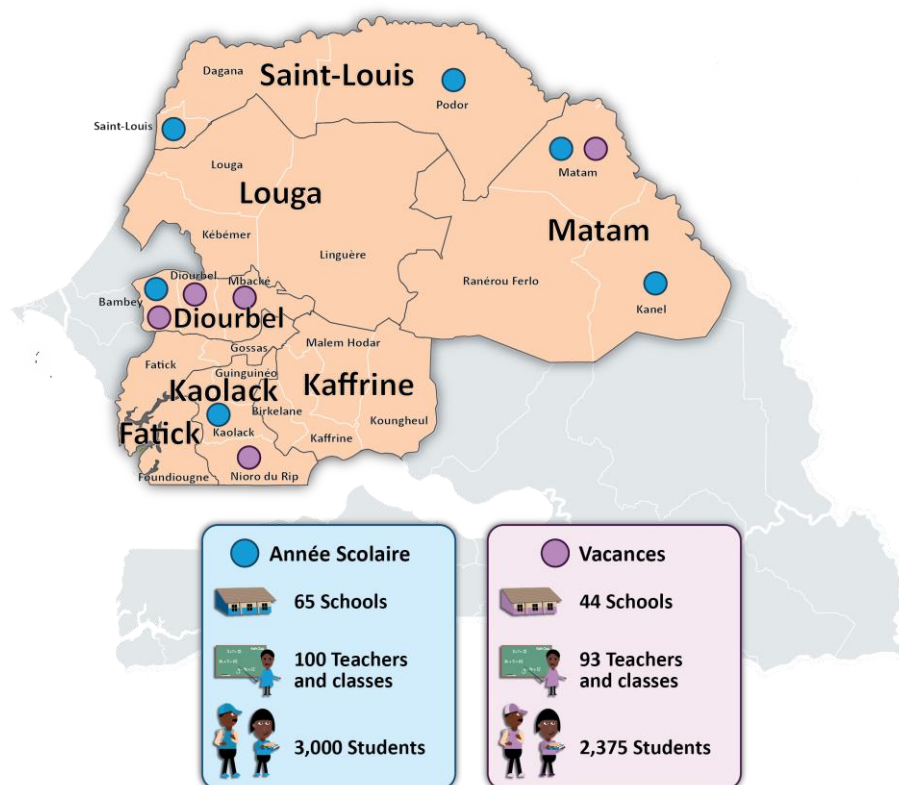
En plus de son appui aux programmes ministériels tels que le MOHEBS et la Remédiation Harmonisée, ARED a commencé, en 2021, à développer son propre programme de remédiation à plus petite échelle, *Ndaw Wune* (« Réussite pour tous »). *Ndaw Wune* est une intervention structurée de rattrapage visant à renforcer les compétences fondamentales en lecture-écriture et en mathématiques chez les élèves sénégalais des premières années du primaire en difficulté, exposés au risque d'échec scolaire et de décrochage, en utilisant l'une des trois principales langues nationales : le pulaar, le sérère ou le wolof. Le programme s'adresse principalement aux élèves de CE1 (*cours élémentaire première année*, ou 3e année) et inclut certains élèves de CP (*cours préparatoire* ou 2^e année du cycle primaire). À l'aide du test présenté à l'**annexe A**, le programme divise les élèves en trois groupes pour la lecture-écriture : un groupe pour les élèves capables d'identifier moins de 10 lettres ; (ii) un groupe pour ceux capables d'identifier au moins 10 lettres mais moins de 10 syllabes ; et (iii) un groupe pour ceux capables de décoder au moins 10 syllabes mais moins de 10 mots. Il comprend deux groupes pour les mathématiques : un groupe pour les élèves capables de réussir correctement moins de cinq comparaisons de nombres et un groupe pour les élèves capables de réussir correctement les cinq comparaisons de nombres (tout en étant éligibles au programme sur la base de leur score en lecture). Les élèves participent à un groupe de lecture-écriture et à un groupe de mathématiques pendant le temps d'enseignement alloué à chaque matière. Les groupes reçoivent un enseignement direct une fois par semaine pour la lecture-écriture et 1,5 fois par semaine pour les mathématiques, tout en participant à des travaux indépendants guidés par l'enseignant à l'aide de cahiers d'exercices adaptés à leur niveau pendant le reste du temps alloué au programme.

Le programme comporte actuellement deux versions : (a) la version Année Scolaire (objet du présent rapport) et (b) la version Vacances. Pour la version *Ndaw Wune* Année Scolaire—qui est la première itération du programme, ARED sélectionne, à l'aide du test décrit à l'Annexe A, les élèves les moins performants de CE1 et de CP pour prendre part à des cours périscolaires de deux heures par jour, trois jours par semaine (soit six heures par semaine au total), pour une durée aussi proche que possible de 25 semaines (durée officielle de l'année

scolaire), compte tenu des contraintes liées au retard de rentrée scolaire, aux vacances, aux conditions météorologiques, aux événements locaux et aux absences des enseignants. Cette approche prévoit un enseignement dans la langue nationale (wolof, pulaar ou sérère) utilisée par les élèves de l'école, dans des classes de 30 élèves. Dans le système éducatif sénégalais, les structures régionales sont appelées IA (*Inspection d'Académie*), et les régions sont divisées en zones desservies par des *Inspections d'Éducation et de Formation* (IEF). Pour l'année scolaire 2024-2025, ARED a mis en œuvre le programme *Ndaw Wune Année Scolaire* dans une centaine de classes comptant au total 3 000 élèves dans six IEF situées dans quatre IA : Diourbel, Kaolack, Matam et Saint-Louis.

La version *Ndaw Wune Vacances* s'appuie sur un modèle similaire, mais est dispensée à raison de trois heures par jour, trois jours par semaine, pendant huit semaines, au cours des longues vacances scolaires (de juin à octobre). Le tableau 2 présente les principales caractéristiques de chacune des versions du programme pour l'année scolaire 2024-2025 et les vacances 2024. La liste complète des endroits où le programme est proposé figure à l'annexe B.

2 . Aperçu du programme *Ndaw Wune* en 2024-2025



À ce jour, aucune évaluation indépendante n'a établi l'impact causal du programme **Ndaw Wune** sur les résultats des élèves en **lecture-écriture** et en **mathématiques**, bien que plusieurs évaluations internes et évaluations de processus aient été réalisées. **Par conséquent, pour produire des données probantes essentielles pour orienter la mise en œuvre des programmes d'apprentissage fondamental en langues nationales et leur passage à l'échelle, l'American Institutes for Research (AIR) a mené une évaluation d'impact et une analyse coût-efficacité de Ndaw Wune Année Scolaire.** Dans le présent rapport, nous exposons les méthodes de notre évaluation et de nos analyses, présentons les résultats de l'étude, ainsi que nos recommandations et conclusions.

2. Théorie du changement

Le modèle des Fondements cognitifs de l'apprentissage de la lecture (*Cognitive Foundations of Reading Acquisition*, CFRA) (Hoover & Tunmer, 2020), qui s'inscrit dans le prolongement de la vision simple de la lecture (Gough & Hoover, 1990), constitue le principal cadre théorique de notre étude. Ce modèle postule que l'acquisition des bases de la compréhension en lecture repose sur deux composantes nécessaires : les capacités de reconnaissance des mots et la compréhension du langage. Quelle que soit la langue dans laquelle un enfant apprend d'abord à lire, il doit à la fois être capable de sonoriser (décoder) les mots avec précision et rapidité, et de comprendre le sens des mots écrits. Les capacités de reconnaissance des mots se composent d'une série de sous-compétences préalables, notamment la connaissance des lettres (à la fois la capacité à nommer les lettres et l'aptitude automatisée à identifier les correspondances graphème-phonème) ainsi que les compétences de décodage (c'est-à-dire la capacité à sonoriser les mots rapidement et avec exactitude).

La compréhension du langage repose également sur une série de sous-compétences préalables : la conscience phonologique (c'est-à-dire la capacité à identifier, reconnaître et manipuler les segments sonores [phonologiques] dans le langage oral), le vocabulaire oral, la syntaxe, ainsi que les connaissances de base nécessaires à la compréhension du langage oral. Il est donc évident que l'acquisition de la compréhension écrite est impossible dans une langue que l'enfant ne comprend pas. Ce constat est non seulement intuitivement fondé, mais également solidement étayé par les travaux empiriques (Ouellette & Beers, 2010).

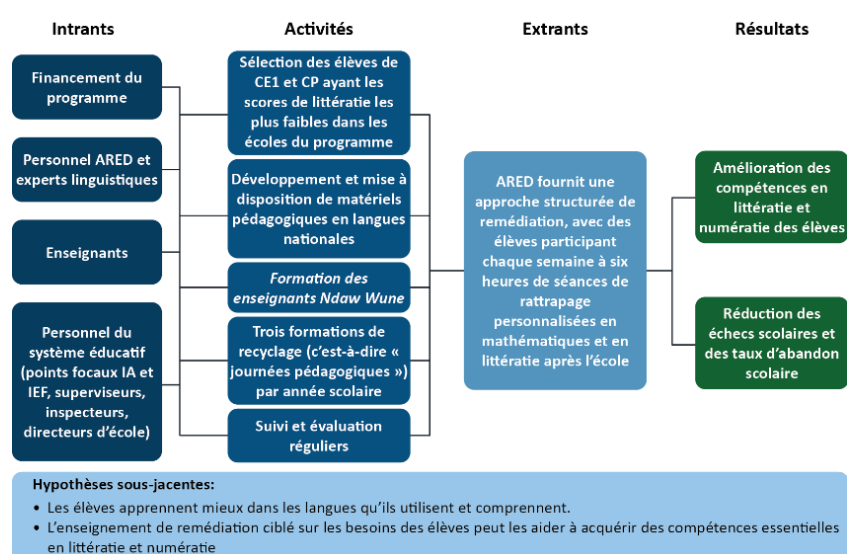
Il est important de noter que l'efficacité de l'enseignement de la lecture est liée à l'existence d'un curriculum, de modèles de formation des enseignants et de supports d'enseignement et d'apprentissage alignés sur ces cadres théoriques (Hoover & Tunmer, 2020). Autrement dit, des programmes, supports et pratiques pédagogiques qui enseignent explicitement la reconnaissance des mots et qui instruisent les enfants dans la/les langue(s) qu'ils comprennent et parlent bien ont une probabilité nettement plus élevée de développer des compétences solides en lecture-écriture que des dispositifs qui ne répondent pas efficacement à ces exigences. Or, au Sénégal, au moins 60 % (et, selon certaines estimations, plus de 73 %) des enfants sur l'ensemble du territoire reçoivent un enseignement dans une langue qu'ils ne comprennent pas (Nakamura & Holla, 2024 ; Nakamura et al., 2024).

Autrement dit, les curricula, supports et pratiques pédagogiques qui enseignent explicitement la reconnaissance des mots et qui dispensent l'enseignement dans la ou les langue(s) que les enfants comprennent et maîtrisent à l'oral ont une probabilité nettement plus élevée de développer des compétences en lecture-écriture que ceux qui ne le font pas de manière effective. Or, au Sénégal, au moins 60 % (et, selon certaines estimations, plus de 73 %) des enfants à l'échelle nationale reçoivent un enseignement dans une langue qu'ils ne comprennent pas (Nakamura & Holla, 2024 ; Nakamura et al., 2024).

En nous appuyant sur cette base théorique pour enseigner la lecture aux enfants dans une **langue qu'ils connaissent**, nous proposons une théorie du changement comme outil précieux pour comprendre comment *Ndaw Wune Année Scolaire*, qui met l'accent sur une approche structurée de la remédiation, contribue à améliorer les résultats d'apprentissage fondamentaux.⁹ Dans le tableau 3, nous présentons une théorie du changement fondamentale pour *Ndaw Wune Année Scolaire*. Dans ce cas, compte tenu du délai, nous nous concentrons sur le résultat immédiat de l'amélioration des compétences en lecture-écriture et mathématiques des élèves, plutôt que sur les effets à moyen terme de la réduction du redoublement et du décrochage scolaire ou sur les effets à long terme de l'amélioration des résultats scolaires sur les élèves et leurs communautés.

L'approche structurée de rattrapage scolaire de *Ndaw Wune* repose sur deux hypothèses clés : (a) les élèves apprennent mieux dans les langues qu'ils utilisent et comprennent, et (b) un enseignement de rattrapage ciblé sur les besoins des élèves en matière de compétences peut les aider à acquérir des compétences essentielles en lecture-écriture et mathématiques. De nombreuses preuves ont démontré les avantages de l'enseignement dans les langues que les élèves utilisent et comprennent (Nag et al., 2019 ; Nakamura et al., 2024 ; Nishanthi, 2020), notamment de meilleurs résultats scolaires dans les langues familières aux enfants et dans les langues internationales ciblées (Evans & Acosta, 2021 ; Koda, 2008 ; Mohohlwane et al., 2024 ; Stern et al., 2024). De même, il a été démontré que l'enseignement de rattrapage, notamment grâce à des méthodes d'apprentissage structurées et à un temps d'enseignement supplémentaire, améliore les résultats scolaires des élèves (Moussa et al., 2021 ; Piper & Dubeck, 2024 ; Snilstveit et al., 2015). ARED a conçu le modèle *Ndaw Wune* en s'appuyant sur ces points forts théoriques et les preuves existantes.

3 . Théorie du changement pour *Ndaw Wune Année Scolaire*



⁹ Nous ne nous appuyons pas sur la lecture en langue seconde, étant donné que le programme vise à favoriser l'alphabétisation des enfants dans leur langue maternelle, à savoir les langues nationales du Sénégal.

Remarque : la présente étude examine uniquement les résultats scolaires des élèves en lecture-écriture et mathématiques comme principaux résultats et ne se penche pas sur l'échec scolaire et les taux d'abandon.

3. Contexte

Le gouvernement sénégalais investit massivement dans le secteur de l'éducation, y consacrant environ 23 % de son budget national et 6 % de son produit intérieur brut (PIB) en 2023 (UNESCO, 2024). Malgré ces engagements financiers importants, des défis persistent : 24 % des enfants en âge d'aller à l'école primaire ne sont pas scolarisés et 69 % des élèves du primaire ne possèdent pas les compétences de base en lecture (Banque mondiale, 2024). En outre, seuls 18 % des élèves âgés de 9 à 16 ans ont un niveau de lecture suffisant, tandis que 22 % seulement ont un niveau suffisant en mathématiques (LARTES-IFAN, 2017). Selon le rapport PASEC 2019, seuls 30 % des élèves sénégalais ont atteint le niveau 4 du PASEC, défini comme la capacité à comprendre des textes narratifs et informatifs, à interpréter des significations implicites et à comprendre les intentions de l'auteur. Des facteurs tels que la qualité des enseignants, le coût de la scolarité et la disponibilité des fournitures scolaires ont été identifiés comme étant corrélés aux taux d'abandon scolaire (Ndour, 2024).

En 2017, dans le but d'améliorer les résultats en lecture des élèves dès les premières années d'école, le Sénégal a mis en œuvre le programme *Lecture Pour Tous* (LPT), financé par l'USAID. À la fin de 2021, le programme avait touché plus de 600 000 élèves de la 1^{re} à la 3^e année dans 3 700 écoles primaires publiques de six régions (Bagby et al., 2022). Le programme LPT a été remplacé par le programme « *Renforcement de la Lecture Initiale pour Tous* » (RELIT), un programme d'enseignement des arts du langage dans les premières années du primaire, qui s'inscrivait dans le cadre de la réforme nationale de l'éducation bilingue au Sénégal jusqu'à sa fermeture, en même temps que l'USAID, au début de l'année 2025. Dans le cadre du RELIT, le déploiement à grande échelle du volet linguistique du programme national d'éducation bilingue grâce à l'aide gouvernementale (G2G) de l'USAID a été étendu d'une région, où l'aide GtG avait été testée avec succès dans le cadre du LPT (Saint-Louis), à six autres régions : Diourbel, Fatick, Kaolack, Kaffrine, Louga et Matam (USAID, 2022). La mise en œuvre directe par RTI, le partenaire de mise en œuvre du RELIT, a été menée dans deux régions, Kédougou et Tambacounda. Bien que le LPT se soit révélé efficace, la plupart des élèves ne lisaient pas couramment en troisième année, ce qui indiquait que des approches de remédiation structurées pourraient être nécessaires pour compléter l'enseignement en classe.

Dans ce contexte, au Sénégal et dans d'autres pays à revenu faible ou intermédiaire (PRFI), les programmes d'enseignement de rattrapage offrent une approche complémentaire pour améliorer l'enseignement dans les classes principales et se sont révélés particulièrement efficaces pour combler les lacunes fondamentales en matière d'apprentissage dans les premières années scolaires (Singh et al., 2024 ; Moussa et al., 2021 ; Brown et al., 2022 ; Duflo et al., 2020). Des programmes tels que « *Teaching at the Right Level* » (TaRL) de Pratham ont permis d'améliorer considérablement la fluidité et la compréhension en lecture chez les élèves du primaire, et des interventions structurées axées sur le sens des nombres, le comptage et les opérations de base ont conduit à des améliorations mesurables des compétences en mathématiques (Duflo et al., 2024 ; Angrist et al., 2023 ; Maruyama et al., 2022). Une revue systématique à grande échelle réalisée par l'International Initiative for Impact Evaluation (3ie) a analysé quatre programmes d'enseignement de rattrapage au Chili,

en Inde et au Mexique, et a constaté une amélioration constante des résultats d'apprentissage, en particulier lorsqu'ils étaient associés à une pédagogie structurée et à un temps d'enseignement supplémentaire (Snilstveit et al., 2016). Les quatre programmes ciblaient des élèves identifiés comme étant en retard par rapport à leurs camarades et proposaient un enseignement sur mesure dans les compétences de base, généralement à des groupes plus petits que les classes habituelles. Les enseignants des quatre programmes étaient soit des bénévoles, soit des personnes recrutées au sein de la communauté locale, et n'étaient pas des enseignants formés.

Ndaw Wune combine les éléments clés d'une pédagogie structurée : un enseignement qui s'appuie sur un champ lexical et des séquences soigneusement conçus pour les fondements théoriques essentiels de l'enseignement de la lecture, des plans de cours variés et des cahiers d'exercices pour les élèves. La littérature spécialisée a démontré que ces éléments, combinés, améliorent l'apprentissage des élèves (Piper et al., 2021). Il existe de nombreuses preuves à l'appui des programmes de pédagogie structurée dans les PRFI, souvent définis comme des groupes d'apprentissage nivelés, fournissant aux enseignants des leçons scriptées pour chaque jour du programme (avec une réduction du script au fur et à mesure que les semaines avancent) et des activités séquencées correspondantes pour les élèves dans des cahiers d'exercices. Ces programmes proposent également généralement une formation et un accompagnement des enseignants alignés sur les principes précédents. Plusieurs études ont été menées sur des programmes pédagogiques structurés à petite échelle, et plus récemment, des preuves ont commencé à émerger concernant des projets à plus grande échelle (Piper & Dubeck, 2024). Par exemple, dans une méta-analyse réalisée en 2020, Evans et Acosta ont constaté que la pédagogie structurée produisait des effets positifs sur les compétences en lecture-écriture et mathématiques dans la plupart des études analysées en Afrique subsaharienne. Une autre méta-analyse a conclu que les programmes de pédagogie structurée avaient « les effets positifs moyens les plus importants et les plus constants sur les résultats d'apprentissage », avec un effet global de 0,23 écart-type (Snilstveit et al., 2017). Une méta-analyse supplémentaire menée dans des pays à revenu faible et intermédiaire (PRFI) a montré que les programmes structurés généraient un effet global de 0,30 écart-type sur les acquis fondamentaux en lecture-écriture, avec des effets plus marqués sur les sous-compétences de reconnaissance lexicale de bas niveau, et des variations d'impact selon le degré d'accompagnement des enseignants et selon que le programme était dispensé en langue première (L1) ou en langue seconde (L2) (Kim, Lee et Zuilkowski, 2020). Une étude examinant les thèmes communs aux programmes mis en œuvre avec succès dans des PRFI a révélé que six des huit programmes relevaient de l'approche structurée en **lecture-écriture** et comprenaient généralement des supports pédagogiques d'apprentissage structuré destinés aux élèves, des guides pédagogiques pour les enseignants, ainsi qu'une formation ciblée assortie d'un accompagnement continu (Stern et al., 2021). En 2023, le Global Education Evidence Advisory Panel a classé la pédagogie structurée parmi les interventions présentant un excellent rapport coût-efficacité (*great buy*).

Cependant, bien que les données probantes relatives à l'impact des programmes de pédagogie structurée soient abondantes, les études évaluant **les effets causaux** et le **rapport coût-efficacité** de programmes de remédiation **déployés à grande échelle** et intégrant une approche de pédagogie structurée en Afrique subsaharienne demeurent rares. Par ailleurs, les éléments probants concernant les effets des programmes de remédiation mis en œuvre par des enseignants formés, pendant le temps scolaire ou en dehors du temps scolaire, sont également très limités. Au Niger, une étude portant sur un programme de remédiation au primaire centré sur la mise à disposition de manuels de mathématiques a mis en évidence un effet de 0,36 écart-type chez les élèves du CE2 au CM2 (Maruyama et Kurosaki, 2021). En Afrique du Sud, un essai contrôlé randomisé évaluant l'impact d'un programme de remédiation destiné aux élèves de du CE2 au CM2 apprenant l'anglais comme langue seconde a observé des effets nuls pour l'ensemble de l'échantillon, mais des effets modérément positifs chez les élèves présentant, au départ, un meilleur niveau en anglais plus élevé (0,12 écart-type) (Fleisch et al., 2017).

Dans ce contexte, en s'appuyant sur les cadres théoriques relatifs au développement de la lecture en langue première à l'échelle internationale et sur les données probantes concernant les programmes de pédagogie structurée dans les PRFI, la présente évaluation examinera les effets causaux d'un programme structuré de remédiation des apprentissages fondamentaux au Sénégal et contribuera ainsi à enrichir la base de données probantes sur l'efficacité des programmes de rattrapage destinés aux élèves en difficulté dont les compétences se situent nettement en deçà du niveau scolaire dans des contextes éducatifs à faibles ressources au sein des PRFI.

4. Questions de recherche

L'évaluation du programme *Ndaw Wune* par AIR aborde les questions de recherche clés suivantes :

- QR1. Dans quelle mesure le programme *Ndaw Wune Année Scolaire* a-t-il un impact sur les résultats en lecture-écriture et mathématiques des élèves de CP et CE1 y participant ?
- QR2. Quel est le coût global et le coût par élève ?
 - QR2A. Comment les coûts varient-ils selon les composantes du programme (par exemple, formation des enseignants, prestation de l'enseignement) ?
 - QR2B. Comment les coûts varient-ils selon le type de dépense (par exemple, personnel, matériel) ?
 - QR2C. Les coûts du programme présentent-ils des variations significatives selon la localisation géographique et, le cas échéant, quelles sont les estimations de coûts par région ?
- QR3. Comment les coûts varient-ils en fonction du scénario de mise à l'échelle du programme (par exemple, coûts liés à l'extension à 50 000 élèves) et quelles économies potentielles peuvent être associées à la modification des calendriers de formation et/ou à la réduction des dépenses de personnel) ?

- QR4. Si la version *Année Scolaire* du programme est efficace, quel est le coût par unité d'effet ?

Étant donné que la théorie du CFRA ainsi que de nombreuses études ont clairement établi que les élèves ne peuvent pas apprendre à lire dans une langue qu'ils ne comprennent pas (en supposant que la lecture soit définie comme la compréhension et non comme le simple décodage, qui repose également fortement sur la connaissance en langage oral), nous formulons l'hypothèse que *Ndaw Wune Année Scolaire* améliorera les sous-compétences fondamentales en lecture-écriture et mathématiques des élèves, dans la mesure où l'enseignement est dispensé dans une langue qu'ils maîtrisent, au moyen d'une pédagogie structurée, de groupes d'apprentissage homogènes par niveau, d'effectifs réduits et de cahiers d'exercices adaptés à leur niveau.

L'évaluation d'impact apporte une réponse à la Question 1, tandis que l'analyse des coûts répond aux Questions 2 et 3. L'évaluation d'impact et l'analyse des coûts contribuent conjointement à répondre à la Question 4, sous la forme d'une analyse coût-efficacité. Nous présentons ci-après les méthodes mobilisées pour chacune des composantes de l'évaluation.

5. Méthodes

5.1. Évaluation d'impact

Nous avons mobilisé la méthode des **doubles différences** (DID) combinée à une approche d'appariement pour analyser l'impact de *Ndaw Wune Année Scolaire* sur les compétences en lecture-écriture et mathématiques des élèves de CP et CE1 participants. Notre méthodologie spécifique est décrite dans la présente section.

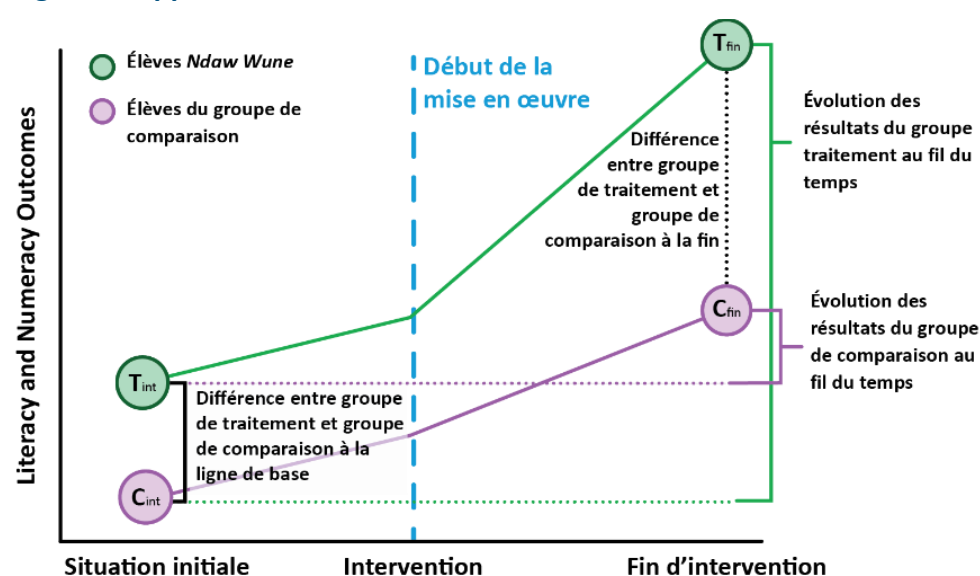
Conception

Nous recourons à une approche de doubles différences combinée à une méthode d'appariement pour estimer l'impact du programme un an après sa mise en œuvre. La figure 4 illustre cette démarche. Plus précisément, nous avons analysé les effets de *Ndaw Wune Année Scolaire* pour la cohorte de l'année scolaire 2024–2025, dont l'enseignement a débuté en décembre 2024 et s'est déroulé sur environ 25 semaines. Cette démarche a consisté, dans un premier temps, à constituer—au moyen de la procédure d'appariement décrite en détail ci-dessous—des groupes d'écoles de comparaison statistiquement similaires (Étape 1), puis à sélectionner des élèves de comparaison selon des modalités analogues à celles retenues pour les élèves du groupe traité (Étape 2). En construisant des groupes statistiquement comparables au niveau des écoles, puis en tenant compte à la fois des écarts initiaux entre les deux groupes au début du programme et de l'évolution des compétences des élèves au cours du temps dans un scénario de « fonctionnement habituel »—tel que reflété par le groupe de comparaison—nous estimons les impacts causaux du *programme Ndaw Wune Année Scolaire*.

Étape 1 : Appariement au niveau scolaire pour identifier les écoles de comparaison. ARED a sélectionné 65 écoles bénéficiaires du programme Ndaw Wune pour l'année scolaire 2024–2025 dans quatre régions (Diourbel, Kaolack, Matam et Saint-Louis). Dans les régions où se trouvaient les écoles sélectionnées, nous avons choisi des « écoles de comparaison » en les appariant en fonction de leur proximité géographique et de leurs caractéristiques scolaires. ARED nous a transmis une liste de 22 539 écoles candidates potentielles au statut d'école de comparaison. Nous avons procédé à une cartographie géographique de ces établissements ;

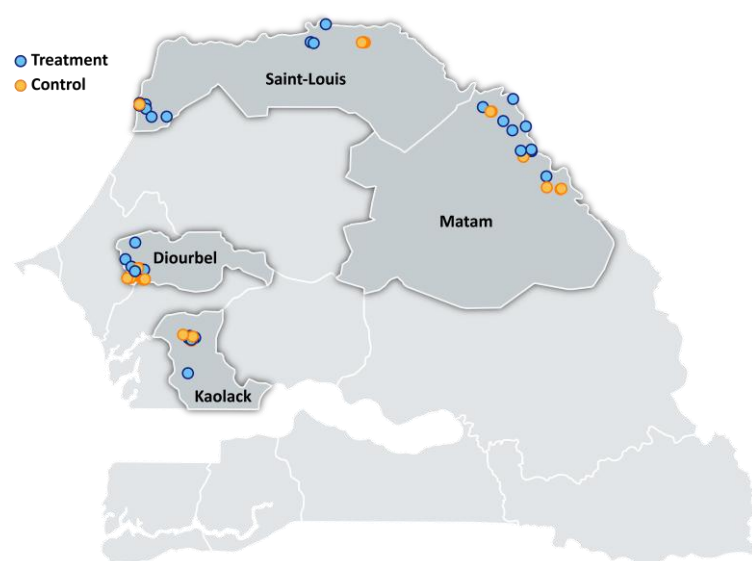
parmi eux, 4 438 se situaient dans les mêmes régions (Diourbel, Kaolack, Matam et Saint-Louis) que les écoles bénéficiaires. En outre, nous avons reçu, pour l'ensemble des écoles du groupe traité et du groupe de comparaison, des données au niveau établissement relatives au département, à la classification urbain/rural, ainsi qu'au nombre total de garçons scolarisés en 2023. Sur la base de ces informations, nous avons d'abord estimé un score de propension (c'est-à-dire la probabilité qu'une école appartienne au groupe traité, sous réserve de l'ensemble des caractéristiques observables) à l'aide d'une régression logit. Les écoles ont ensuite été appariées en minimisant la différence entre leurs probabilités prédites, chaque école traitée étant appariée à l'école de comparaison la plus similaire à l'aide d'un appariement par plus proche voisin. L'appariement a abouti à 35 paires d'écoles de traitement/comparaison qui étaient les plus proches, pour un total de 70 écoles (voir le **tableau 5** pour l'emplacement des écoles de traitement et de comparaison).¹⁰

Figure4 . Approche DID



¹⁰ Sur la base des hypothèses suivantes : avec un échantillon de 70 écoles (35 écoles du groupe de traitement et 35 écoles du groupe témoin) et 16 élèves par école, nous serions en mesure de détecter un effet de 0,22 écart-type (ET), en supposant un niveau de puissance de 80 % ; une corrélation intra-cluster de 0,15 pour le niveau 2 (écoles) ; et un R au carré de 0,40 pour le niveau 1 (élèves) et de 0,40 pour le niveau 2 (écoles). L'étude n'avait pour objectif que de détecter les effets sur l'ensemble de l'échantillon et non sur différents sous-groupes. Nous ne fournissons donc que des preuves suggestives des effets sur les sous-groupes. Les calculs ont été effectués à l'aide de l'outil PowerUp! L'ampleur des effets des programmes de rattrapage varie entre 0,06 et 0,27 écart-type pour les résultats en mathématiques et en lecture (Álvarez Marinelli et al., 2024 ; Jacob & Lefgren, 2002 ; Özek, 2021).

5 . Emplacement des écoles de traitement et de comparaison



Bien que les données disponibles au niveau des établissements pour procéder à l'appariement des écoles de traitement et de comparaison aient été restreintes, comme le montrent **les tableaux 6 et 7**, l'appariement a permis d'identifier des écoles qui étaient pour la plupart statistiquement similaires au départ. Plus précisément, **l'annexe 6** examine l'équilibre des variables qui ont été incluses dans l'appariement, et **l'annexe 7** examine l'équilibre ex post au niveau des écoles sur la base des données de référence qui ont été collectées. Dans notre analyse finale, nous intégrons comme variables de contrôle celles qui présentaient des écarts statistiquement significatifs à la ligne de base.

Tableau 6. Équilibre au niveau des écoles (variables appariées)

Variable	Moyenne des écoles de comparaison	Moyenne des écoles de traitement	Test t apparié Valeur p
Bambey	14	34	0,052*
Kanel	31	26	0,603
Kaolack	3	9	0,310
Matam	23	11	0,210
Podor	9	11	0,695
Saint-Louis	20	9	0,177
Urbain	34	34	0,803
Nombre de garçons scolarisés	115	98	0,359
Nombre de filles scolarisées	14	34	0,052

Remarque : Les erreurs standards sont indiquées entre parenthèses. $N = 35$ écoles du groupe de comparaison et 35 écoles du groupe de traitement.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau 7. Équivalence de base au niveau scolaire (variables collectées au début de l'étude)

Variable	Moyenne des écoles de comparaison	Moyenne des écoles de traitement	Test t apparié Valeur p
Électricité	86	83	0,747
Eau potable	94	86	0,238
Latrines	83	66	0,104
Station de lavage	37	14	0,029
Savon	20	3	0,025
Repas gratuits	11	3	0,168

Remarque. Les erreurs standards sont indiquées entre parenthèses. $N = 35$ écoles de traitement et 35 écoles témoins, à l'exception d'une école de traitement où la question relative au savon n'a pas pu être observée.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Étape 2 : Sélection des élèves du groupe de traitement et du groupe de comparaison. Après avoir établi la liste des écoles appariées du groupe de traitement et du groupe de comparaison, AIR a sélectionné un groupe d'élèves du groupe de comparaison pour imiter la sélection des élèves dans les écoles du groupe de traitement. Pour garantir la comparabilité, nous avons reproduit l'approche utilisée par ARED pour sélectionner les élèves éligibles au programme.

Dans les écoles participant au programme, ARED a fait passer une évaluation de base en lecture-écriture et mathématiques aux élèves de CE1 (voir exemple à l'**annexe A**). Les élèves ayant obtenu un score inférieur à un seuil prédéterminé (par exemple, la capacité de lire au moins 10 mots) ont été considérés comme éligibles au programme. Lorsque plus de 35 élèves de CE1 remplissaient les critères d'éligibilité, les élèves ayant obtenu les scores les plus faibles étaient retenus en priorité. Lorsque moins de 35 élèves de CE1 étaient éligibles, ARED administrait la même évaluation aux élèves de CP afin de pourvoir les places restantes. Une fois la liste définitive des bénéficiaires établie par ARED, celle-ci nous a été transmise et nous avons procédé à un tirage aléatoire de 16 élèves par école participant au programme pour constituer l'échantillon d'élèves participant au programme.

Pour s'assurer que les élèves de comparaison inclus dans notre échantillon présentent des caractéristiques aussi proches que possible de celles des élèves sélectionnés pour bénéficier d'un rattrapage dans les écoles de *Ndaw Wune*, nous avons reproduit la procédure mise en œuvre par ARED et administré le même test de sélection aux élèves des écoles de comparaison, en appliquant le même principe de priorisation des élèves de CE1. Après avoir identifié, dans chaque établissement, les 35 élèves éligibles selon les critères retenus, nous avons procédé à un tirage aléatoire de 16 élèves pour constituer l'échantillon des élèves du groupe de comparaison. Le **tableau D1** de l'**annexe D** décrit la répartition des notes obtenues par les élèves des écoles de traitement et de comparaison lors du test de sélection.¹¹

¹¹ Il convient de noter que, compte tenu de la nature très spécifique de la sélection des élèves pour ce programme, cette étude n'est représentative que de la population échantillonnée pour *Ndaw Wune* (principalement des élèves de CE1).

Étape 3. Collecte des données de référence et finales pour l'analyse en doubles différences.

Une fois les élèves du groupe de comparaison et du groupe de traitement identifiés à l'Étape 2, nous avons collecté des données quantitatives de référence au début de l'année scolaire (novembre-décembre 2024) et des données de suivi à la fin de l'année scolaire (mai-juin 2025). En mobilisant l'approche des **doubles différences**, nous avons comparé l'évolution moyenne des résultats dans le temps pour le groupe traité à l'évolution moyenne observée pour le groupe de comparaison. Cette approche a permis d'identifier les impacts du programme en isolant, d'une part, les différences initiales entre les deux groupes au début de l'étude, ainsi que les tendances générales dans les domaines étudiés (capturées par le groupe de comparaison) — ces dernières étant des changements survenus en raison de facteurs sans rapport avec le programme étudié, ces dernières correspondant à des changements indépendants du programme évalué, tels que l'influence d'interventions concomitantes. L'analyse en doubles différences a estimé les effets de *Ndaw Wune* sur les résultats des élèves en comparant les variations avant/après des indicateurs d'apprentissage entre les élèves des groupes de traitement et de comparaison.

Deux caractéristiques essentielles de l'estimateur en **doubles différences** le rendent particulièrement pertinent pour l'identification d'effets causaux non biaisés. Premièrement, l'utilisation de mesures pré- et post-intervention permet de « différencier » les caractéristiques individuelles et scolaires non observées mais fixes (c'est-à-dire invariantes dans le temps) susceptibles d'influencer les résultats, telles que le genre ou la localisation de l'établissement. Cette approche nous permet également de comparer l'évolution de l'indicateur à sa valeur en l'absence d'intervention. Deuxièmement, l'utilisation de l'évolution observée dans le groupe de comparaison comme **contrefactuel** nous permet de tenir compte des tendances générales de la valeur du résultat. L'hypothèse d'identification fondamentale est celle de la **tendance parallèle**, selon laquelle les trajectoires des résultats dans les écoles de traitement et de comparaison auraient évolué de manière similaire en l'absence du programme.¹² Nous estimons cette hypothèse plausible dans la mesure où les écoles de traitement et de comparaison présentaient des caractéristiques largement similaires avant le début du programme (voir les **tableaux 6 et 7** ci-dessus).

Analyse

Nous avons estimé l'impact du programme *Ndaw Wune* au moyen d'un modèle de régression en **doubles différences**, en intégrant des variables de contrôle relatives aux caractéristiques individuelles des élèves. L'échantillon a été analysé à partir de l'équation suivante:

$$Y_{ist} = \alpha + \beta_1 Treatment_s + \beta_2 Post_t + \beta_3 (Treatment_s * Post_t) + \delta Z_{ist} + \varepsilon_{ist}$$

Où :

- Y_{ist} est le résultat recherché (par exemple, la compréhension écrite) de l'élève i dans l'école s au moment t ;

¹² Compte tenu de la courte durée de l'évaluation (à savoir une année scolaire), nous n'avons pas été en mesure de tester l'hypothèse des tendances parallèles. En d'autres termes, l'approche DID partait du principe qu'il n'existait aucune tendance variable dans le temps entre le groupe traité et le groupe témoin. Pour évaluer l'hypothèse des tendances parallèles, il aurait fallu disposer de données couvrant plusieurs périodes antérieures au traitement.

- $Treatment_s$ est une variable indicatrice égale à 0 pour les écoles du groupe de comparaison (sans intervention) et égale à 1 pour les écoles du groupe de traitement (*Ndaw Wune*) ;
- $Post_t$ est égal à 0 pour la base de référence (avant l'intervention) et à 1 pour la ligne de fin (après l'intervention) ;
- Z_{ist} comprend un ensemble de variables individuelles au niveau des élèves qui sont incluses comme variables de contrôle (note, variables indicatrices pour les filles, les élèves bénéficiant d'un soutien scolaire externe et la langue dans laquelle l'enquête a été menée) ; et
- ε_{ist} est le terme d'erreur.

Le dispositif en doubles différences permet d'estimer l'effet en intention de traiter (*Intent-to-Treat, ITT*) du programme *Ndaw Wune* (β_3), c'est-à-dire l'effet moyen du traitement pour les élèves ou les écoles assignés à une condition de traitement, indépendamment de l'adhésion au traitement. Les analyses ont été conduites sur l'échantillon complet, puis désagrégées selon le **genre**, la **région** et la **langue d'instruction**. Nous avons également réalisé des tests de robustesse en intégrant des variables de contrôle additionnelles au niveau de l'école et de l'élève, lorsque celles-ci présentaient des différences statistiquement significatives entre les groupes traité et de comparaison au départ.

Taille des échantillons et calendrier de collecte des données

Notre équipe a recueilli des données quantitatives de référence auprès des élèves du groupe de traitement et du groupe de comparaison entre le 27 novembre et le 1^{er} décembre 2024, juste avant le début de la mise en œuvre du programme en décembre 2024. Les données de suivi permettant de mesurer les impacts sur une année scolaire¹³ ont été recueillies entre le 8 mai et le 10 juin 2025. Le plan initial consistait à suivre les mêmes élèves du début à la fin dans le cadre d'une étude longitudinale. Au début, dans 70 écoles avec un objectif de 1 120 élèves, nous avons inclus 1 109 élèves dans l'échantillon (561 élèves du groupe de comparaison et 548 élèves du groupe de traitement). Les élèves de chaque école de traitement et de comparaison ont été déterminés sur la base de l'étape 2 mentionnée précédemment (c'est-à-dire les élèves dont les résultats étaient inférieurs au seuil lors du test de sélection initial). Cependant, à la fin de l'étude, nous n'avons pu interroger que 77 % de l'échantillon initial. À la fin de l'étude, pour les 23 % d'élèves que nous n'avons pas pu suivre depuis le début, des élèves de remplacement ont été sélectionnés pour maintenir une taille d'échantillon final similaire à celle de l'échantillon initial. Ces élèves de remplacement ont été sélectionnés à partir des tests de sélection initiaux pour garantir que seuls les élèves ayant obtenu un score inférieur au seuil initial soient retenus comme remplaçants.¹⁴ Dans l'analyse finale, nous avons effectué des tests de robustesse à la fois sur l'échantillon apparié (élèves présents à la fois lors de l'évaluation initiale et de l'évaluation finale) et sur l'échantillon

¹³ Dans ce contexte, « une année scolaire » équivaut à environ 6-7 mois (ou 22-28 semaines).

¹⁴ Si les élèves qui se situaient en dessous du seuil de sélection au début de l'étude n'étaient pas disponibles comme remplaçants le jour de la collecte des données, le test de sélection était refait et ceux qui se situaient en dessous du seuil étaient sélectionnés comme remplaçants. Comme c'était la fin de l'année scolaire, il est possible que les élèves sélectionnés comme remplaçants sur la base du test de sélection effectué à la fin de l'étude soient différents de ceux sélectionnés au début, car une année scolaire entière d'apprentissage s'était écoulée. Nous examinons l'attrition différentielle et effectuons également une analyse de robustesse avec uniquement l'échantillon apparié (élèves à la fois au début et à la fin de l'étude) et les échantillons complets (élèves d'origine plus élèves de remplacement).

complet (élèves présents à la fois lors de l'évaluation initiale et finale et élèves de remplacement). Le **tableau 8** présente les tailles d'échantillon pour les évaluations initiale et finale, y compris les échantillons manquants et de remplacement.

Le taux d'attrition global de l'échantillon était de 23 %, celui du groupe de traitement étant de 25 % et celui du groupe de comparaison de 20 %, soit une attrition différentielle de 5 %. Cette combinaison d'attrition globale et différentielle demeure en deçà des seuils prudent et optimiste établis par le *What Works Clearinghouse* pour caractériser une attrition différentielle élevée.¹⁵ Dans la section 6, nous présentons l'équivalence de base pour l'échantillon complet, l'échantillon apparié et les élèves perdus de vue par rapport aux élèves appariés. Nous effectuons également une analyse de robustesse avec uniquement l'échantillon apparié (élèves à la fois au début et à la fin) et les échantillons complets (élèves d'origine plus élèves de remplacement). Dans l'analyse de l'échantillon complet, les échantillons sont traités comme des sections transversales répétées (c'est-à-dire qu'aucune imputation des valeurs de base n'est effectuée pour l'échantillon de remplacement).

8 . Taille des échantillons d'élèves

Groupe	Échantillon cible	Échantillon de référence réel	Échantillon final apparié	Étudiants de remplacement	Échantillon final complet
Comparaison (35 écoles)	560	561	446	111	557
Traitement (35 écoles)	560	548	412	149	561
Total	1 120	1 109	858	260	1 118

Mesures des résultats

Nous avons recueilli des données quantitatives à partir (a) d'enquêtes auprès des élèves et (b) d'évaluations des compétences en lecture-écriture et mathématiques des élèves.

Enquête auprès des élèves. AIR a mené une enquête auprès des élèves et recueilli des informations sur des sujets clés, notamment des informations générales et des informations sur la langue parlée à la maison et l'environnement scolaire. Nous nous sommes appuyés sur divers cadres théoriques décrivant le lien entre le climat scolaire et le développement des élèves, tels que la théorie des systèmes écologiques (Bronfenbrenner, 2000), qui postule que le développement de l'apprentissage des élèves est influencé par le contexte de la classe, de l'école et de la communauté. Les enquêteurs ont mené cette enquête oralement, en utilisant un langage adapté aux enfants dans la langue nationale que les enfants utilisent et comprennent bien.

Évaluations de la lecture-écriture et des mathématiques. Pour mesurer les compétences fondamentales en **lecture-écriture** dans les langues nationales utilisées comme langues d'enseignement, nous avons mobilisé un instrument d'évaluation fondé sur notre cadre théorique des apprentissages fondamentaux, ainsi que sur des outils validés sur le plan psychométrique et adaptés au contexte local, déjà utilisés au Sénégal par AIR et d'autres

¹⁵ Les limites prudentes et optimistes de l'attrition différentielle pour un taux d'attrition global de 23 % sont respectivement de 5,1 % et 9,5 % (*What Works Clearinghouse™ Standards Handbook Version 4.0*, n.d.).

partenaires. Ces outils ont permis d'évaluer les compétences phonologiques des enfants, leur compréhension du langage (y compris le vocabulaire réceptif et expressif et la compréhension orale), leur connaissance des symboles/lettres (son et nom, selon le cas), leur décodage de mots réels et non réels, leur fluidité en lecture à voix haute et leurs compétences en compréhension écrite. Pour évaluer les compétences fondamentales en mathématiques, nous avons utilisé une version adaptée de l'évaluation des mathématiques dans les premières années d'école (EGMA) pour évaluer les compétences des élèves en matière de comptage oral, d'identification des nombres, de discrimination des quantités, de nombres manquants et d'addition et de soustraction. Le **tableau 9** présente les concepts couverts par les évaluations en lecture-écriture et en mathématiques, et une traduction des évaluations complètes est présentée à **l'annexe F**. Tous les tests ont été réalisés dans la langue d'enseignement, qui était la langue nationale sénégalaise choisie comme langue officielle de l'école.

Tableau 9 . Évaluation des compétences en lecture-écriture et mathématiques

	Concept	Sous-tâche (si différente du concept)	Description
Lecture-écriture	Compréhension orale		L'enfant écoute une courte histoire, qui lui est lue deux fois, et répond aux questions de compréhension correspondantes.
	Dominance linguistique (vocabulaire expressif)	Fluidité sémantique	L'enfant nomme autant de mots que possible dans toutes les langues qu'il prétend connaître (même un peu) en une minute, en utilisant une image d'une scène familière comme aide, puis répète l'activité dans d'autres langues qu'il connaît.
	Conscience phonologique expressive		L'enfant écoute des mots familiers qui lui sont lus à haute voix, puis identifie le son initial (phonème) de chaque mot.
	Connaissance des sons des lettres		L'enfant voit une liste de lettres (y compris certains digrammes, diphtongues et combinaisons de lettres) et produit le son représenté par la lettre ou le groupe de lettres.
	Connaissance de la correspondance graphème-phonème (décodage)	Décodage de mots familiers	L'enfant prononce à haute voix une liste de mots couramment utilisés dans la langue ciblée.
		Décodage de pseudo-mots	L'enfant prononce à voix haute une liste de pseudo-mots qui suivent les règles orthographiques de la langue cible. ¹⁶
	Fluidité de lecture à voix haute		L'enfant lit à haute voix une courte histoire et l'enquêteur note le nombre correct de mots prononcés avec rapidité et précision en 60 secondes.
	Compréhension écrite	Compréhension écrite	L'enfant répond à des questions de compréhension liées à la nouvelle de la sous-tâche de fluidité de lecture à voix haute (uniquement pour les sections qu'il a lues).

¹⁶ Cette sous-tâche permet de déterminer si un enfant est capable de décoder des mots familiers à vue, plutôt que de les connaître.

	Concept	Sous-tâche (si différente du concept)	Description
mathématiques	Identification des nombres		L'enfant identifie oralement les nombres écrits sous leur forme standard, en passant des nombres à un chiffre aux nombres à trois chiffres.
	Comptage		L'élève regarde des images comportant des objets répétés et compte combien il en voit.
	Discrimination des nombres		L'enfant se voit présenter une paire de nombres et doit identifier lequel des deux est le plus grand, avec une progression graduelle allant des nombres à un chiffre aux nombres à trois chiffres.
	Sens des nombres	Exercice sur les nombres manquants	L'enfant se voit présenter une suite de quatre nombres suivant un schéma régulier, dont l'un est manquant ; il doit identifier le nombre manquant.
	Addition		L'enfant répond à des exercices d'addition dont le niveau de difficulté augmente progressivement.
	Soustraction		L'enfant répond à des exercices de soustraction dont le niveau de difficulté augmente progressivement.

Collecte de données

Avant la collecte des données de référence, AIR a obtenu l'autorisation de son comité d'éthique interne. AIR a également étudié la possibilité de faire appel à des comités d'éthique locaux, mais a conclu que cela n'était pas nécessaire, car l'étude ne collectait pas de données médicales. L'équipe chargée de la collecte des données, composée de 16 enquêteurs et de quatre superviseurs, a suivi une formation en présentiel à Dakar, animée par Dalberg et à laquelle AIR a participé à distance. Des parties prenantes d'ARED ont également assisté à la formation et apporté leur précieuse contribution. Les enquêteurs ont été recrutés spécifiquement en fonction de la langue d'enseignement utilisée dans chaque région. Avant le recrutement, Dalberg savait clairement quelle langue serait utilisée dans chaque école et chaque commune, ce qui lui a permis de sélectionner des enquêteurs qui non seulement parlaient ces langues, mais maîtrisaient également les dialectes locaux utilisés dans les communes sélectionnées. Cette maîtrise des dialectes était essentielle pour garantir une communication claire et une compréhension totale entre les enquêteurs et les élèves, pour assurer un référentiel commun sur chacune des sous-tâches évaluées entre les élèves et les enquêteurs. De plus, tous les enquêteurs recrutés étaient bilingues ou multilingues, chacun parlant au moins deux langues en plus du français, notamment le wolof et le sérère ou le pulaar. Cette polyvalence linguistique a été un facteur clé pour maintenir une collecte de données de haute qualité dans des contextes linguistiques diversifiés.

La formation, qui s'est déroulée du 20 au 22 novembre 2024, comprenait une journée d'essai des outils dans une école non participante au programme à Dakar, les enquêteurs s'exerçant à utiliser les outils au moins une fois. La collecte des données de référence s'est déroulée du 27 novembre au 21 décembre. Les superviseurs de Dalberg ont contacté les directeurs d'école en amont afin de planifier les visites et de minimiser les perturbations des activités scolaires. Les données ont été saisies sur tablettes via le logiciel Tangerine. La collecte des données de fin de période a suivi les mêmes procédures que la collecte des données de référence et s'est

déroulée du 8 mai au 10 juin. À la fin, 13 des 16 enquêteurs initiaux ont participé de nouveau à la collecte, complétés par trois remplaçants. Une formation de remise à niveau de deux jours a eu lieu les 5 et 6 mai 2025.

Après avoir reçu les jeux de données de référence et de fin de période nettoyées de Dalberg, nous avons effectué des contrôles de qualité pour garantir l'exactitude des données. Plus précisément, nous avons utilisé Stata pour rechercher d'éventuelles valeurs aberrantes, des réponses incohérentes, des observations dupliquées, entre autres. Toute préoccupation (par exemple, des doublons potentiels, des élèves codés comme ayant passé l'enquête en français, ou des règles d'arrêt appliquées de manière inconsistante) a été transmise à l'équipe de collecte et résolue avec l'appui de leurs responsables de collecte. Une fois le jeu de données jugé satisfaisant, nous avons calculé les variables de résultats pour chacun des sous-tests de lecture-écriture et de mathématiques, puis standardisé les scores. En plus des contrôles de qualité, AIR a effectué une analyse de fiabilité inter-évaluateurs (IRR) pour vérifier la cohérence entre les enquêteurs, qui a donné des résultats positifs (voir **annexe E**). Pour ce faire, plusieurs enquêtes (15 % au début et 20 % à la fin) ont été menées avec un intervenant et un observateur, l'intervenant administrant l'enquête comme d'habitude et l'observateur observant et enregistrant ses propres réponses sur sa tablette sans participer au questionnaire. Leurs réponses ont été comparées à l'aide d'une régression par analyse de variance (ANOVA). L'analyse a mis en évidence un niveau élevé de cohérence entre enquêteurs pour chaque résultat vérifié.

ARED a signalé certains items des sous-tests comme étant trop difficiles, mais pour maintenir la comparabilité entre les données de référence et les données fin de période, nous avons décidé de ne pas modifier le test. Nous avons plutôt effectué une analyse psychométrique pour vérifier la robustesse du test. L'analyse n'a pas révélé de relation forte entre les items signalés par ARED et ceux identifiés par l'analyse psychométrique. Plus précisément, pour évaluer la qualité des items du test, nous avons mené une série d'analyses psychométriques à l'aide de la théorie classique des tests (CTT)¹⁷ et de la modélisation de Rasch.¹⁸ Ces méthodes nous aident à identifier les items qui pourraient ne pas donner de bons résultats et qui devraient être exclus de la notation. Nous avons réalisé la modélisation de Rasch de deux manières : l'une incluant tous les items dans toutes les langues (étalonnage simultané) et l'autre examinant les items séparément par langue. Chaque approche offre des informations complémentaires sur la difficulté et les performances des items. Nous avons effectué la modélisation de Rasch avec un calibrage simultané à l'aide du package TAM dans R pour estimer la difficulté des items dans toutes les langues au sein de chaque domaine (par exemple, la compréhension orale). Cette approche nous permet de comparer la difficulté des

¹⁷ Selon la théorie classique des tests, la difficulté d'un item est définie comme la proportion de répondants qui ont répondu correctement à l'item. Des valeurs plus élevées indiquent des items plus faciles, tandis que des valeurs plus faibles suggèrent une plus grande difficulté. Idéalement, les items devraient se situer dans une fourchette de difficulté comprise entre 0,2 et 0,8. Les items dont les valeurs sont inférieures à 0,1 ou supérieures à 0,9 sont signalés comme potentiellement trop difficiles ou trop faciles, respectivement, et sont examinés en vue de leur exclusion.

¹⁸ La modélisation de Rasch est une approche statistique utilisée pour estimer la difficulté d'un item et la capacité d'une personne sur une échelle commune, où la difficulté de l'item reflète le point auquel une personne a 50 % de chances de répondre correctement. Les scores de difficulté sont exprimés sous forme de scores z, où les valeurs négatives indiquent des items plus faciles, zéro représente une difficulté moyenne et les valeurs positives indiquent des items plus difficiles. Idéalement, les scores devraient se situer entre -2 et +2.

items entre les langues au sein d'un même domaine. Par exemple, un item avec un score de difficulté de -1 en pulaar est plus facile qu'un item avec un score de 0 en sérère. Cependant, les comparaisons entre différents domaines (par exemple, écoute vs lecture) ne sont pas valables. Nous avons également effectué des analyses Rasch par langue et par domaine à l'aide du package eRm sous R. Bien que cette approche ne permette pas de comparer directement la difficulté des items entre les langues ou les domaines, elle fournit des statistiques d'ajustement précieuses qui nous aident à évaluer si les items se comportent comme prévu.¹⁹ L'évaluation a démontré de solides performances psychométriques dans toutes les langues, ce qui confirme son adéquation pour l'évaluation d'impact. Les résultats convergents de la théorie classique des tests et de l'analyse Rasch indiquent que les modules sont bien ciblés. La difficulté des items se situe dans la fourchette moyenne souhaitée, la difficulté Rasch est modérée et les valeurs moyennes quadratiques Infit se situent dans les limites recommandées de 0,7 à 1,3. Des tailles d'échantillon robustes dans ces domaines renforcent en outre la stabilité des estimations, tandis que les modules de vocabulaire illustré et de compréhension présentent des effectifs substantiels, avec seulement quelques écarts isolés. Pris dans leur ensemble, ces éléments étayent la validité de la structure interne et la fiabilité satisfaisante des scores en pulaar, sérère et wolof.

Limites de l'évaluation d'impact

L'évaluation d'impact a rencontré quelques limites. Premièrement, les données disponibles au niveau des établissements pour l'appariement des écoles de traitement et de comparaison étaient limitées. Idéalement, l'appariement aurait pu s'appuyer sur des caractéristiques observables, notamment les taux d'assiduité des élèves et des enseignants, les taux d'abandon scolaire, l'existence de cantines scolaires et d'autres ressources éducatives et, surtout, la langue d'enseignement. Cependant, la plupart de ces données n'étaient pas disponibles avant la collecte des données lors de l'appariement. Ainsi, bien que l'échantillon apparié au niveau des écoles repose sur le nombre limité de variables disponibles, les appariements pourraient différer selon de nombreuses autres caractéristiques, observables comme non observable. Deuxièmement, notre puissance statistique pour détecter les effets hétérogènes était limitée. Bien que la taille de notre échantillon ait été suffisante pour évaluer les effets globaux, notre conception de l'évaluation inclut un nombre limité d'écoles, ce qui limite notre capacité à détecter les hétérogénéités dans les estimations d'impact avec une précision suffisante. En outre, une hypothèse clé lors de l'utilisation de la méthodologie des différences dans les différences est que les élèves du groupe de traitement et du groupe de comparaison suivaient des parcours identiques avant le début de la mise en œuvre (c'est-à-dire que sans *Ndaw Wune*, les élèves du groupe de traitement et du groupe de comparaison auraient obtenu les mêmes résultats).

Cela est généralement vérifié en collectant des données à plusieurs moments avant le début de la mise en œuvre. Il est courant (comme dans le cas présent) que la collecte de données antérieures ne soit pas possible, et la stratégie d'appariement et les tests d'équivalence de base permettent de pallier cette limitation. Enfin, nous n'avons pas accès aux informations

¹⁹ Plus précisément, nous avons examiné deux types de statistiques d'ajustement : *l'infit* et *l'outfit*. L'infit est sensible aux schémas de réponse inattendus sur les items qui correspondent au niveau de capacité d'une personne, tandis que l'outfit est plus sensible aux réponses inhabituelles sur les items qui sont très faciles ou très difficiles pour une personne. Ces statistiques nous aident à déterminer si les items fonctionnent correctement, par exemple si les items plus difficiles sont systématiquement répondu correctement par les personnes ayant des capacités supérieures.

sur la langue d'enseignement des écoles avant l'échantillonnage. Au final, nous disposions de données pour 12 écoles sérères du groupe de traitement et une seule école sérère du groupe de comparaison. Cet échantillon n'était pas suffisamment large pour permettre des comparaisons fiables entre les groupes de traitement et de comparaison. L'échantillon comprenait 12 écoles wolof du groupe de traitement et 23 écoles wolof du groupe de comparaison, ainsi que 11 écoles pulaar du groupe de traitement et du groupe de comparaison.

5.2. Analyse des données de suivi

Pour compléter l'évaluation, ARED a partagé un sous-ensemble de données issues de ses exercices de suivi continu. Ces données comprenaient (a) le nombre d'absences des élèves entre décembre 2024 et mars 2025 et (b) de brefs commentaires de 54 enseignants (répartis entre 10 superviseurs dans les quatre régions cibles) concernant le programme entre décembre 2024 et avril 2025 (avec 552 observations au total). Les chercheurs d'AIR ont utilisé les commentaires des enseignants pour fournir des informations supplémentaires sur le fonctionnement du programme. Cependant, les données sur l'absentéisme des élèves comportaient plusieurs incohérences,²⁰ ce qui nous a empêchés de tirer des conclusions fiables.

Méthode d'analyse

Notre équipe a suivi un processus systématique pour analyser qualitativement les commentaires des enseignants. À l'aide du logiciel de recherche qualitative NVivo, nous avons d'abord téléchargé l'ensemble des commentaires et classé les données par (a) superviseur, (b) région et (c) école. Nous avons ensuite codé les commentaires de manière inductive en fonction des thèmes émergents.

Limites des données de suivi

Notre utilisation des données de suivi d'ARED dans cette évaluation est soumise aux limites suivantes :

Couverture partielle de la période de mise en œuvre. Nous n'avons reçu aucun commentaire des enseignants après avril 2025, nous n'avons donc pas pu inclure les données du dernier mois de mise en œuvre.

Brièveté des commentaires des enseignants. Les commentaires transmis par les enseignants se limitaient à de brèves formulations portant sur la mise en œuvre du programme, telles que « difficulté à comprendre certains exercices » ou « appréciation des parents ». Bien que ces données apportent des informations utiles sur les succès et les défis rencontrés par les enseignants et les élèves, elles ne permettent pas d'approfondir l'analyse ni de fournir des éléments de détail, par exemple sur la nature précise des exercices que les enseignants et les élèves ont eu du mal à comprendre.

5.3. Analyse des coûts

Pour répondre aux questions d'évaluation 2 à 4, AIR a mené une analyse rétrospective des coûts du programme de rattrapage scolaire *Ndaw Wune*. Nous avons donc recueilli les coûts

²⁰ Les incohérences comprenaient des rapports manquants provenant des journées d'enseignement de *Ndaw Wune* et des rapports manquants provenant des tuteurs et des écoles pour les journées incluses.

du point de vue du programme, y compris tous les coûts directement engagés par le responsable de la mise en œuvre du programme. Notre équipe a produit des estimations de coûts pour la version 2024-2025 du programme : *Ndaw Wune Année Scolaire*.

Horizon temporel. Nous avons basé notre analyse des coûts sur les programmes mis en œuvre au cours de l'année scolaire 2024-2025. Nous avons utilisé les coûts liés aux régions et au nombre d'élèves et d'écoles desservis par le programme pendant cette période. Pour mieux refléter ce qui serait considéré comme une mise en œuvre « normale » du programme, nous avons procédé à plusieurs ajustements des coûts. Par exemple, au cours de l'année scolaire 2024-2025, ARED a adapté ses procédures habituelles d'évaluation (c'est-à-dire qu'elle n'a réalisé qu'une évaluation de mi-parcours, avec une taille d'échantillon ajustée) afin de s'aligner sur l'évaluation d'impact conduite par AIR et d'éviter les redondances. En conséquence, nous avons estimé les coûts liés à l'évaluation programmatique à partir de procédures susceptibles d'être appliquées au cours d'une année scolaire standard (à savoir **deux vagues de collecte de données**), en retenant les localisations des écoles et les niveaux de prix observés dans le cadre du programme en 2024–2025.²¹

Approche. AIR a adopté une approche hybride pour l'analyse des coûts, intégrant à la fois des méthodes de calcul des coûts basées sur les intrants et sur les activités.

L'approche par intrants constitue une méthodologie de référence, certes exigeante en ressources, et largement reconnue comme une bonne pratique (Danks & Kendziora, 2021 ; Levin et al., 2018). Elle repose sur un processus systématique visant à identifier et quantifier l'ensemble des intrants du programme (par exemple, le nombre d'heures-personnes par rôle et par activité, la nature et la quantité des matériels utilisés, etc.), puis à attribuer un prix unitaire à chaque intrant afin de calculer son coût. L'approche par intrants permet une décomposition très précise des coûts par activités constitutives. Toutefois, elle est fortement consommatrice de données et requiert une contribution substantielle de l'ensemble des personnels et participants au programme pour recueillir des informations détaillées sur les quantités et les types spécifiques de ressources utilisées. Pour optimiser le rapport coût-efficacité et de minimiser la charge de travail liée à la collecte de données pour le personnel et les participants, nous avons limité notre collecte de données pour l'étude des coûts aux documents programmatiques et financiers qu'ARED nous a fournis et aux entretiens avec les principaux membres du personnel d'ARED. À partir des données fournies, nous avons suivi l'approche par intrants pour les ressources pour lesquelles nous avons pu obtenir des données détaillées sur les prix unitaires et les quantités, telles que le nombre de brochures destinées aux élèves et les frais de déplacement remboursés aux enseignants pendant la formation.

Pour les intrants du programme pour lesquels nous ne disposons pas de données au niveau « intrant », nous avons complété l'analyse par une approche fondée sur les activités. L'approche par activités consiste à identifier, pour chaque activité du programme, les intrants nécessaires, puis à en affecter la valeur à partir des budgets du programme, des pièces de dépenses, des registres d'activités et d'entretiens avec des informateurs clés (Brander et al., 2024). Autrement dit, au lieu de recueillir des données sur le nombre d'heures consacrées par

²¹ Le G50 fait référence au *Groupe élargi de concertation et de coordination des partenaires au développement du Sénégal*, un groupe reconnu au niveau national qui normalise les taux de remboursement des frais de déplacement et les indemnités versées au personnel gouvernemental et aux collaborateurs des projets.

chaque membre du personnel d'ARED à chaque activité dans le cadre du programme *Ndaw Wune*, puis de multiplier ce nombre par le coût horaire de la main-d'œuvre pour déterminer les coûts de personnel pour chaque composante du programme, l'approche basée sur les activités s'appuie sur les descriptions des activités, les données de rémunération annuelle et des entretiens avec le personnel d'ARED pour répartir proportionnellement un coût de personnel estimé entre les différentes activités du programme. Plus précisément, la comptabilité par activité a été utilisée pour ventiler les coûts de personnel d'ARED entre les quatre composantes du programme et pour estimer les coûts de déplacement lorsque des itinéraires étaient disponibles sans pièces de dépense associées. Cette approche hybride constitue une méthodologie efficiente permettant de produire des données de coûts rigoureuses et opérationnelles, susceptibles de servir de base à la mise à l'échelle du programme.

Sources des données

Pour obtenir les informations nécessaires à l'analyse des coûts, notre équipe (a) a reçu d'ARED des données sur le budget et les dépenses, tant au niveau du programme qu'au niveau organisationnel et (b) a conduit des entretiens, organisé des discussions approfondies et correspondu par e-mail avec le personnel concerné du programme. Nos échanges avec ARED ont notamment porté sur les sujets suivants :

- les composantes du programme ainsi que les tâches et activités associées ;
- le personnel mobilisé, leurs niveaux de rémunération et le temps consacré au programme ;
- les infrastructures, équipements et matériels utilisés dans le cadre du programme, ainsi que leurs coûts ;
- les ressources affectées à l'administration du programme.

Les données que nous avons reçues d'ARED comprenaient notamment :

- les informations salariales du personnel d'ARED ainsi que la proportion de leur consacrée à *Ndaw Wune* ;
- les coûts d'impression et de livraison des supports pédagogiques ;
- le nombre d'écoles, d'élèves et d'enseignants concernés, ainsi que la localisation des établissements participant à *Ndaw Wune Année Scolaire* ;
- la liste des activités prévues dans le cadre de *Ndaw Wune Année Scolaire* ;
- les termes de référence (TDR), les itinéraires et/ou les budgets pour les formations des enseignants, les journées pédagogiques et les évaluations ;
- le protocole gouvernemental pour le remboursement des frais de déplacement (protocole G50).

À partir des estimations fournies par les informateurs et des données sur les dépenses fournies par ARED, nous avons reconstitué de manière rétrospective les intrants mobilisés et les quantités correspondantes pour la mise en œuvre de *Ndaw Wune Année Scolaire* en 2024–2025. Lorsque des données sur le coût des intrants faisaient défaut, nous avons formulé des

hypothèses pour estimer au mieux les coûts du programme. Par exemple, nous avons reçu les termes de référence et l'itinéraire d'une série de journées pédagogiques (mais aucun document budgétaire), et nous avons utilisé le budget de déplacement d'une série d'évaluations, complété par des conversations avec ARED, pour estimer les frais de déplacement liés aux journées pédagogiques. Nous avons documenté toutes les hypothèses que nous avons formulées (voir **annexe C**), que nous avons partagées avec le personnel chargé du programme et des finances d'ARED afin qu'il les vérifie et/ou les corrige par courrier électronique et lors de deux réunions virtuelles.

Analyse

L'équipe d'analyse des coûts d'AIR a enregistré les données relatives aux coûts dans une base de données appelée « modèle de coûts des ressources » (MCR). Dans le MCR, nous avons répertorié les coûts pour chacune des activités suivantes du programme :

- **Formation initiale des enseignants** : la formation des enseignants de 4 jours qui a lieu au début de chaque cycle du programme, précédée d'une « journée d'harmonisation » d'une journée pour les formateurs.
- **Enseignement direct** : enseignement hebdomadaire dispensé à *Ndaw Wune*, comprenant notamment le coût des supports pédagogiques et la rémunération des enseignants.
- **Journées pédagogiques** : formation de remise à niveau d'une demi-journée, précédée d'observations en classe, organisées à deux reprises au cours de l'année scolaire.
- **Gestion, planification, suivi et évaluation** : tous les autres coûts liés à la mise en œuvre du programme, tels que le temps consacré aux tâches administratives liées à la fourniture de matériel didactique, à la préparation de la formation et au suivi et à l'évaluation continus. Cette catégorie comprend également les coûts liés aux deux séries d'évaluations des élèves qu'ARED organise lorsque les ressources le permettent.

Nous avons réparti les coûts en cinq catégories, en conservant chaque ressource comme un élément distinct au sein de chaque catégorie :

- **Personnel** : tout le temps consacré au programme par le personnel, y compris le personnel du programme ARED, les consultants chargés d'animer les formations et les indemnités versées aux enseignants, aux inspecteurs et aux superviseurs.
- **Supports pédagogiques et équipement** : supports (par exemple, livres pour les enseignants et les élèves) et équipement (par exemple, ordinateurs portables et projecteurs) mobilisés pour le fonctionnement et la mise en œuvre du programme.
- **Déplacements** : toutes les dépenses liées au transport et à l'hébergement, telles que les remboursements de frais kilométriques, les frais de chauffeur et de carburant, l'hébergement et les repas/indemnités journalières.
- **Installations** : tous les coûts liés à l'utilisation d'espaces physiques pour le programme. Il convient de noter que l'utilisation de salles de classe pour l'enseignement direct, la formation des enseignants et les journées pédagogiques n'a entraîné aucun coût pour le programme et a donc été exclue. Nous prenons en compte les coûts de location

d'espaces pour la formation des évaluateurs. Les coûts liés aux espaces de bureau sont inclus dans la catégorie des frais généraux, décrite ci-dessous.

- **Frais généraux** : coûts tels que les locaux à usage de bureaux, les fournitures et autres dépenses générales liées au fonctionnement de l'organisation. ARED budgétise ces dépenses en allouant 10 % des coûts du projet aux frais généraux, ce que nous avons reflété dans l'analyse des coûts.

Nous avons ensuite attribué à chaque élément son prix et sa quantité. Le RCM calcule ensuite les coûts de chaque intrant, en les résumant par catégorie de ressources et par composante du programme, et présente les résultats en F CFA et en USD. À partir de ces estimations de coûts globaux, nous avons divisé par 3 000 pour estimer les coûts par élève.

Bien que nous ayons initialement prévu d'estimer séparément les coûts liés au fonctionnement annuel, à la phase d'installation (pour fournir des formations et des supports pédagogiques dans les nouvelles écoles de la même région) et à la mise à l'échelle (traduction des supports et établissement de nouveaux partenariats), nous avons appris, lors d'un entretien avec le responsable du programme *Ndaw Wune* que, puisque le programme en est encore à la phase de « recherche-action », les supports pédagogiques ont, jusqu'à présent, été révisés chaque année. En conséquence, nous avons intégré les coûts de révision des supports dans notre estimation des coûts, même si le matériel sera stabilisé après les révisions de 2026 et ces coûts ne devraient donc plus être engagés de manière continue.

Limites des données d'analyse des coûts

Comme toute étude, notre analyse des coûts est soumise à certaines limites et certains risques, que nous signalons ci-dessous :

Approche hybride. Bien que notre approche hybride combinant la méthode par intrants et la méthode par activités, soit à la fois pertinente et efficace pour les besoins de la présente étude, elle implique un arbitrage en termes de précision. Pour les activités du programme pour lesquelles nous ne disposons pas de données détaillées sur les dépenses, nous avons estimé les coûts sur la base des descriptions des activités, des termes de référence. Les documents programmatiques comportaient un niveau limité de détail concernant la ventilation des coûts de personnel et des dépenses de transport, ce qui nous a conduits à formuler plusieurs hypothèses substantielles pour ces catégories de coûts. Nous avons communiqué nos hypothèses préliminaires à ARED par courrier électronique et lors d'une réunion virtuelle, avons demandé des éclaircissements et des données supplémentaires, et avons modifié nos hypothèses en fonction des commentaires d'ARED. Nous avons ensuite organisé une réunion supplémentaire avec ARED, au cours de laquelle celle-ci a communiqué les coûts précis de personnel devant être imputés à *Année Scolaire*. Toutes les hypothèses sont documentées de manière exhaustive à **l'annexe C**.

Biais de rappel. Les données de coûts recueillies sont susceptibles d'être affectées par un biais de rappel de la part des responsables de la mise en œuvre du programme. La triangulation entre les données de coûts documentées et les informations issues des entretiens a permis d'atténuer les limites associées à ce biais. En cas de divergence entre les données documentaires et les données issues des entretiens, nous avons privilégié les données documentées.

Fluctuation du taux de change. Conformément aux bonnes pratiques, l'analyse des coûts a été réalisée en monnaie locale (FCFA). Nous avons retenu un taux de 580 F CFA pour 1 USD, correspondant à celui intégré dans les fichiers budgétaires fournis par ARED. Toutefois, le taux de change a connu des variations substantielles au cours de la période de mise en œuvre du programme, passant d'environ 642 F CFA pour 1 USD en février 2025 au taux actuel d'environ 563 F CFA pour 1 USD. Les estimations de coûts en USD doivent donc être interprétées avec prudence.

5.4. Analyse coût-efficacité

À partir des résultats issus de l'analyse des coûts et de l'évaluation d'impact, nous avons apprécié les effets observés du *programme Ndaw Wune Année Scolaire* au regard des coûts de mise en œuvre du programme. Dans une analyse coût-efficacité, les chercheurs comparent généralement les coûts du programme à un scénario de « statu quo », c'est-à-dire aux coûts d'opportunité ou aux conditions de comparaison auxquelles les élèves auraient été exposés en l'absence de l'intervention. Dans le cas présent, le scénario « statu quo » correspond à l'absence de programme de rattrapage. Par conséquent, le coût de référence du « statu quo » retenu pour la comparaison est de 0 \$.

Étant donné que l'évaluation d'impact portait sur plusieurs compétences en lecture-écriture et en mathématiques, nous avons retenu une approche d'analyse coût-efficacité mieux adaptée aux programmes à objectifs multiples, souvent qualifiée d'analyse coût-conséquences ou d'inventaire des impacts. Cette approche est particulièrement pertinente pour l'analyse des coûts de *Ndaw Wune*, dans la mesure où les coûts nécessaires pour générer les différents effets en lecture-écriture et en mathématiques ne peuvent pas être dissociés. Il n'est pas possible, par exemple, d'attribuer une proportion donnée des coûts du programme aux seuls impacts en lecture-écriture et une autre proportion aux impacts en mathématiques. Cette approche est particulièrement pertinente pour l'analyse des coûts de *Ndaw Wune*, dans la mesure où les coûts nécessaires pour générer les différents effets en lecture-écriture et en mathématiques ne peuvent pas être dissociés. Il n'est pas possible, par exemple, d'attribuer une proportion donnée des coûts du programme aux seuls impacts en lecture-écriture et une autre proportion aux impacts en mathématiques. Pour faciliter la comparaison avec les coûts et les impacts d'autres programmes de rattrapage scolaire, nous présentons également une transformation linéaire des impacts par élève obtenus pour chaque investissement de 100 dollars.

6. Résultats

6.1. Évaluation de l'impact

Équivalence de référence au niveau des élèves

Nous avons présenté l'équilibre au niveau scolaire dans **les tableaux 4 et 5** ; dans la présente section, nous commençons par examiner l'équivalence de base au niveau des élèves pour l'échantillon complet et uniquement pour l'échantillon apparié. Comme indiqué dans la section 4.1, l'appariement a été effectué au niveau scolaire, ce qui a donné un échantillon relativement équilibré au niveau scolaire. Les élèves des groupes de traitement et de comparaison ont ensuite été sélectionnés sur la base de tests de dépistage en lecture-écriture et mathématiques, ces tests étant administrés de la même manière dans les écoles de traitement et dans les écoles de comparaison appariées. **Le tableau 10** présente les différences de base dans les caractéristiques observables entre les élèves des écoles de traitement et de comparaison sur la base des données de référence complètes. Il existait certaines différences notables entre les élèves des groupes de traitement et de comparaison au départ, la langue étant l'une des plus importantes. Alors que le groupe de traitement était réparti de manière presque égale entre le wolof, le sérère et le pulaar, le groupe de comparaison était composé de seulement 3 % de élèves sérérophones (provenant d'une seule école), de 31 % de locuteurs pulaar et de 66 % d'élèves wolophones. En outre, 88 % des élèves du groupe de traitement ont déclaré bénéficier d'un soutien scolaire externe, contre 76 % du groupe de comparaison ($p < 0,01$).²² Le test de lecture-écriture initial a révélé une différence entre les groupes en matière de connaissance des sons des lettres : le groupe de comparaison a obtenu un résultat supérieur de 4,6 lettres (sur 100 ; $p < 0,10$) à celui du groupe de traitement. Plusieurs différences ont été observées pour les sous-compétences en mathématiques. Le groupe de traitement a obtenu un score supérieur à celui du groupe de comparaison en matière de comptage, de discrimination des chiffres et d'addition ; toutefois, l'ampleur de la différence est assez faible pour chacune de ces compétences, inférieure à une réponse correcte.

Compte tenu des différences observables entre les élèves du groupe de traitement et ceux du groupe de comparaison (malgré la sélection des élèves à l'aide du même test), nous présentons des tests de robustesse, y compris toutes les covariables de référence qui sont statistiquement significatives en tant que contrôles dans l'estimation principale (**tableau D6, annexe D**).

Tableau 10 . Caractéristiques de base au niveau des élèves

Résultat	Groupe de comparaison	Groupe de traitement	Différence
Filles	50	50	0
Âge	9,51 ans	9,64 ans	0,14 an
Note	95	96	1

²² Il est possible que les élèves des écoles participant au projet aient mal compris et considéré *Ndaw Wune* comme un soutien scolaire externe.

Résultat	Groupe de comparaison	Groupe de traitement	Différence
A fréquenté une école maternelle	58	59	2
LE : wolof	66	33	-33 %***
LE : sérère	3	35	32 %***
LE : pulaar	31	32	1
Nourriture consommée aujourd'hui	83	82	-1
A reçu des cours particuliers en dehors de l'école	76	88	12
Au cours de la dernière semaine complète d'école, combien de jours avez-vous manqué ?	1,56 jour	1,38 jour	-0,18 jour**
Compréhension orale	3,14	3,4	0,27
Conscience phonologique	7,59	8,46	0,87
Connaissance des sons des lettres	75,3	70,75	-4,55*
Décodage de pseudo-mots	6,64	5,99	-0,65
Décodage de mots familiers	6,11	6,74	0,64
Fluidité de lecture à voix haute	5,24	5,57	0,33
Compréhension écrite	0,54	0,68	0,14
Identifiant numérique	18,07	18,23	0,16
Comptage	5,11	5,33	0,22**
Discrimination des chiffres	9,74	10,39	0,66
Sens des nombres	4,4	4,48	0,05
Addition	6,21	7,01	0,80
Soustraction	4,76	5,12	0,35

Remarque : LE = langue d'enseignement. L'échantillon comprend l'ensemble des élèves interrogés lors de l'évaluation initiale.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Par ailleurs, comme indiqué précédemment à la Section 5.1, le taux d'attrition global s'élevait à 23 %, avec une attrition différentielle de 5 %. Pour nous assurer qu'il n'y avait aucune raison systématique à l'absence d'élèves à la fin de l'étude, nous avons analysé l'équivalence de base entre les élèves appariés et les élèves perdus de vue (voir **tableau D2, annexe D**). Il n'y avait que quelques différences dans l'échantillon de traitement : l'échantillon manquant comptait davantage de élèves wolophones que l'échantillon apparié (43 % contre 30 %, $p < 0,10$), et l'échantillon apparié comptait davantage de locuteurs pulaar (35 % contre 23 %, $p < 0,05$). Les élèves perdus de vue étaient plus susceptibles de se trouver à Kaolack (37 % contre 20 %, $p < 0,05$) et moins susceptibles de se trouver à Saint-Louis (10 % contre 24 %, $p < 0,05$). Il n'y avait aucune différence à Matam. Les élèves perdus de vue (au début de l'étude) ont également déclaré avoir manqué 0,3 jour d'école de plus la semaine précédente par rapport aux élèves

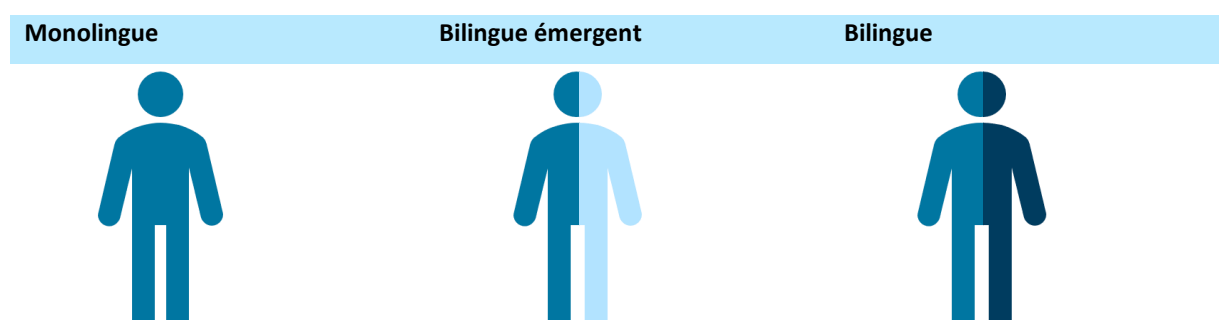
appariés ($p < 0,05$). Aucune des variables de résultat liées à la lecture-écriture ou aux mathématiques ne différait entre les élèves perdus de vue et les élèves appariés.

Au sein du groupe de comparaison, les élèves perdus de vue étaient plus susceptibles d'être des filles (65 %) que les élèves appariés (47 %, $p < .01$). Par ailleurs, les élèves perdus de vue étaient davantage concentrés à Kaolack (44 % contre 29 %, $p < .01$) et moins présents à Saint-Louis (30 % contre 22 %, $p < .05$), sans différence observée pour Matam. S'agissant des scores en lecture-écriture fondamentale, la seule différence relevée est que les élèves appariés étaient, en moyenne, capables de lire 1,5 mot familier de plus (sur 50) que les élèves perdus de vue ($p < .10$). Compte tenu de certaines différences de caractéristiques observables entre les échantillons apparié et manquant, nous avons réalisé, à titre de test de robustesse, les régressions d'impact sur l'échantillon complet et sur l'échantillon apparié ; les résultats se sont révélés similaires. En outre, nous présentons l'équivalence à la ligne de base en ne retenant que l'échantillon apparié dans **le tableau D3, annexe D**, et avec l'échantillon complet sans l'échantillon sère dans **le tableau D4, annexe D**.

Analyse descriptive de la fluidité sémantique

Dans le cadre de la collecte de données de référence et de fin de période, nous avons déterminé de manière empirique le niveau de maîtrise des langues que les élèves déclaraient parler. Nous avons utilisé des évaluations de la fluidité sémantique, qui mesurent le vocabulaire oral expressif, pour déterminer la ou les langues dominantes des élèves dans le cas où un enfant recevait un enseignement dans une langue nationale différente de sa langue maternelle. Sur la base de ces données, les élèves ont été classés en trois catégories : **monolingues**, **bilingues émergents** et **bilingues**. Les élèves monolingues étaient ceux qui déclaraient ne connaître qu'une seule langue., tandis que les élèves bilingues émergents étaient ceux qui déclaraient connaître deux langues et qui obtenaient un écart de plus de cinq mots dans le nombre de mots expressifs qu'ils étaient capables de produire dans chaque langue en 60 secondes. Les élèves bilingues étaient ceux qui étaient capables de produire moins de cinq mots dans chaque langue qu'ils déclaraient connaître. Il s'agissait d'une variable importante pour s'assurer que les élèves étaient effectivement évalués et instruits dans des langues qu'ils comprenaient et parlaient principalement (en tant qu'enfants monolingues ou bilingues). Les définitions utilisées pour classer les enfants dans différentes catégories de compétences linguistiques sont décrites dans **l'annexe 11**.

11



L'enfant affirme ne connaître qu'une seule langue.

L'enfant affirme connaître plus d'une langue, mais a obtenu un écart de plus de cinq points entre les deux langues.

L'enfant affirme connaître plus d'une langue et a obtenu un score inférieur à cinq points dans les deux langues.

Selon le test de fluidité sémantique, la plupart des enfants (93 % au début de l'étude et 94 % à la fin) étaient monolingues dans une langue nationale. Par exemple, au début de l'étude, 87 % des élèves étaient monolingues en wolof, contre 91 % à la fin. Une petite fraction des élèves était bilingue (7 % des élèves échantillonnés au début de l'étude et 5 % à la fin) et moins de 1 % était en train de devenir bilingue. Parmi les élèves bilingues, la proportion d'élèves urbains était plus élevée. La langue la plus connue était le wolof, suivi du pulaar et du sérère (**tableau 12**). Plus précisément, le **tableau 12** montre la répartition des langues que les élèves connaissent dans l'échantillon. Par exemple, 51 % des élèves ont le wolof comme l'une de leurs langues dominantes au début et à la fin de l'étude. La plupart de ces élèves sont monolingues : 87 % des élèves wolophones (51 % de l'échantillon au début et à la fin de l'étude) sont monolingues au début de l'étude et 91 % d'entre eux sont monolingues à la fin de l'étude. De même, 86 % des élèves pulaarophones (34 % de l'échantillon au début et à la fin de l'étude) sont monolingues au début de l'étude et 87 % d'entre eux sont monolingues à la fin de l'étude.

Figure12 . Langue dominante chez les élèves

Langue	% de dominance dans la langue au début de l'étude	N	% dominant dans la langue à la fin de l'étude	N
Wolof	50,77	1 109	51,16	1 118
Pulaar	33,63	1 109	34,26	1 118
Sérère	18,94	1 109	18,52	1 118
Français	1,89	1 109	0,36	1 118
Sonike	1,44	1 109	0,36	1 118
Bambara	0,27	1 109	0	1 118
Bassari	0,09	1 109	0,09	1 118
Diola	0,09	1 109	0	1 118

Presque l'ensemble des élèves maîtrisaient la langue nationale correspondant à la langue d'enseignement de leur établissement. Le taux d'inadéquation entre la langue dominante de l'élève et la langue d'enseignement s'établissait à 2 % à la ligne de base et à 3 % à la fin de période. Le niveau d'inadéquation est faible pour les trois langues d'enseignement : le wolof, le sérère et le pulaar (**tableau 13**).

13 . Adéquation entre la langue dominante des élèves et la langue d'enseignement

Pourcentage d'inadéquation	Fin	N	Référence	N
Wolof	2	558	4	552
Sérère	0	206	2	208
Pulaar	1	354	4	349

Effets globaux du traitement

Le **tableau 14** montre l'impact du *programme Ndaw Wune Année Scolaire* sur les compétences en lecture-écriture et mathématiques mentionnées dans le tableau 9 ci-dessus. Ces effets sont présentés sous forme d'effets standardisés²³ dans la sixième colonne. Chaque cellule de la sixième colonne correspond à une estimation de régression distincte. Les lignes surlignées en vert et accompagnées d'un astérisque correspondent à des résultats positifs et significatifs. Toutes les régressions tiennent compte des éléments suivants : niveau scolaire, sexe, suivi ou non d'un soutien scolaire externe et langue dans laquelle l'enquête a été menée.²⁴ Les résultats sont présentés pour l'échantillon complet (élèves appariés entre le début et la fin de l'étude et élèves de remplacement).

Dans l'ensemble, les données montrent **que Ndaw Wune Année Scolaire a eu un impact positif sur les scores des sous-compétences fondamentales en lecture-écriture et sur les scores de compréhension écrite**. L'impact le plus important a été observé sur le décodage de mots familiers²⁵, suivi de près par la connaissance des sons des lettres. En termes de **décodage de mots familiers**, les élèves pouvaient, en moyenne, décoder six mots familiers sur 50 au début de l'étude, tandis que les élèves du groupe de traitement pouvaient en décoder sept. À la fin de l'étude, les élèves du groupe de traitement ont amélioré leur score de 12 points de pourcentage pour atteindre 13 mots, tandis que les élèves du groupe de comparaison ont amélioré leur score de 6 points de pourcentage pour atteindre 9 mots, ce qui équivaut à un effet du traitement de 0,47 *écart-type* ($p < 0,01$). En ce qui concerne **la connaissance des sons des lettres** au début de l'étude, les élèves du groupe de comparaison pouvaient identifier les sons de 75 lettres sur 100, contre 71 pour les élèves du groupe de traitement. À la fin de l'étude, les élèves du groupe de comparaison et ceux du groupe de traitement ont augmenté leurs scores de connaissance des sons des lettres à 78 et 84 sur 100, respectivement, ce qui équivaut à un effet du traitement de 0,45 *écart-type* ($p < 0,01$). Au début de l'étude, les élèves du groupe de comparaison pouvaient décoder en moyenne sept **pseudo-mots** sur 50, contre six pour les élèves du groupe de traitement. À la fin de l'étude, les élèves du groupe de comparaison ont augmenté leur score à 10, tandis que ceux du groupe de traitement sont passés à 12 mots, ce qui correspond à un effet du traitement de 0,32 *écart-type* ($p < 0,01$). Ces trois compétences de décodage combinées donnent un effet du traitement de 0,42 *écart-type* ($p < 0,01$).

²³ Calculé en soustrayant la moyenne finale de la moyenne initiale et en divisant par l'écart type initial.

²⁴ Nous n'incluons pas les jours d'absence comme contrôle, car ceux-ci ont été déclarés par les élèves eux-mêmes et ne sont donc pas très fiables. Nous les incluons toutefois comme contrôle dans la section consacrée aux contrôles de robustesse.

²⁵ Les « mots familiers » désignent les mots réels qui apparaissent dans le programme scolaire des élèves et qu'ils sont censés être capables de décoder.

Ndaw Wune a également eu un impact faible mais positif et significatif sur les scores **de fluidité en lecture à voix haute** des élèves. Au début de l'étude, les élèves du groupe de comparaison pouvaient lire cinq mots sur 50, contre six pour les élèves du groupe de traitement. À la fin de l'étude, les élèves du groupe de comparaison lisaient 10 mots, tandis que ceux du groupe de traitement avaient amélioré leurs scores pour atteindre une moyenne de 13 mots. L'augmentation plus importante de la fluidité de lecture à voix haute chez les élèves du groupe de traitement s'est traduite par un effet du traitement de 0,32 *écart-type* ($p < 0,01$). Enfin, **la compréhension écrite** des élèves du groupe de traitement s'est également améliorée. Au début de l'étude, les élèves du groupe de comparaison obtenaient en moyenne 0,5 réponse correcte sur six, contre 0,7 pour les élèves du groupe de traitement. À la fin de l'étude, les deux groupes ont amélioré leurs scores, atteignant respectivement 1,0 et 1,4, ce qui correspond à un effet du traitement de 0,14 *écart-type* ($p < 0,10$).

Deux des cinq sous-tests de mathématiques ont montré des effets significatifs de *Ndaw Wune*. Le premier, **le sens des nombres**, demandait aux élèves de compléter la série de nombres manquants, par exemple : 1, __, 3, 4. Ici, les élèves du groupe de traitement sont passés de 4,5 sur 12 au début de l'étude à 5,8 à la fin, tandis que les élèves du groupe de comparaison sont passés de 4,4 à 5,3, ce qui équivaut à un effet du traitement de 0,19 *écart-type* ($p < 0,05$). Le deuxième impact positif concernait le sous-test **de soustraction**, qui posait des questions de soustraction de plus en plus difficiles. Les élèves du groupe de traitement sont passés d'une moyenne de 5,1 sur 12 à 7,3 à la fin de l'étude, tandis que les élèves du groupe de comparaison sont passés de 4,8 au début de l'étude à 6,2 à la fin, soit un effet du traitement de 0,2 *écart-type* ($p < 0,05$).

Nous avons également examiné les effets pour les élèves appariés dans les écoles du groupe de traitement en fonction des groupes de numératie *Ndaw Wune* dans lesquels les élèves avaient été placés au début de l'étude (c'est-à-dire les élèves qui pouvaient effectuer moins de cinq comparaisons de nombres [le groupe des nombres] et les élèves qui pouvaient effectuer les cinq comparaisons de nombres [le groupe des opérations]). Les élèves des deux groupes de mathématiques ont montré une amélioration statistiquement significative des scores moyens pour tous les résultats que nous avons testés. Les élèves du groupe des nombres ont obtenu une amélioration plus importante que le groupe des opérations, plus avancé, dans quatre des six résultats de mathématiques mesurés, y compris les trois résultats de niveau inférieur : identification des nombres, comptage et discrimination des chiffres.

Tableau 14. Effets du traitement (échantillon complet avec témoins)

Sous-tâche	Groupe de comparaison – Situation de référence	Comparaison finale	Référence du traitement	Fin du traitement	Effet standardisé du traitement	N
Compréhension orale (sur 6)	3,14	4,22	3,40	4,41	-0,04 (0,10)	2 227
Conscience phonologique (sur 10)	7,59	7,35	8,46	8,74	0,15 (0,13)	2 227
Test de décodage	s/o	s/o	s/o	s/o	0,42*** (0,08)	2 227

Sous-tâche	Groupe de comparaison – Situation de référence	Comparaison finale	Référence du traitement	Fin du traitement	Effet standardisé du traitement	N
• Connaissance des sons des lettres (sur 100)	75,31	77,69	70,75	83,52	0,45*** (0,10)	2 227
• Décodage de pseudo-mots (sur 50)	6,64	9,94	5,99	11,85	0,32*** (0,08)	2 227
• Décodage de mots familiers (sur 50)	6,11	9,00	6,74	12,87	0,47*** (0,10)	2 227
Fluidité de lecture à voix haute (sur 50)	5,24	9,88	5,57	13,12	0,32*** (0,11)	2 227
Compréhension écrite (sur 6)	0,54	0,97	0,68	1,35	0,14* (0,08)	2 191
Numéro d'identification (sur 24)	18,07	20,34	18,23	21,00	0,11 (0,07)	2 227
Comptage (sur 6)	5,11	5,51	5,33	5,69	-0,03 (0,11)	2 227
Discrimination des chiffres (sur 12)	9,74	10,85	10,39	11,22	-0,10 (0,08)	2 227
Sens des nombres (sur 12)	4,43	5,25	4,48	5,84	0,19** (0,08)	2 227
Ajout (sur 12)	6,21	8,27	7,01	9,28	0,07 (0,08)	2 227
Soustraction (sur 12)	4,76	6,22	5,12	7,26	0,20** (0,09)	2 227

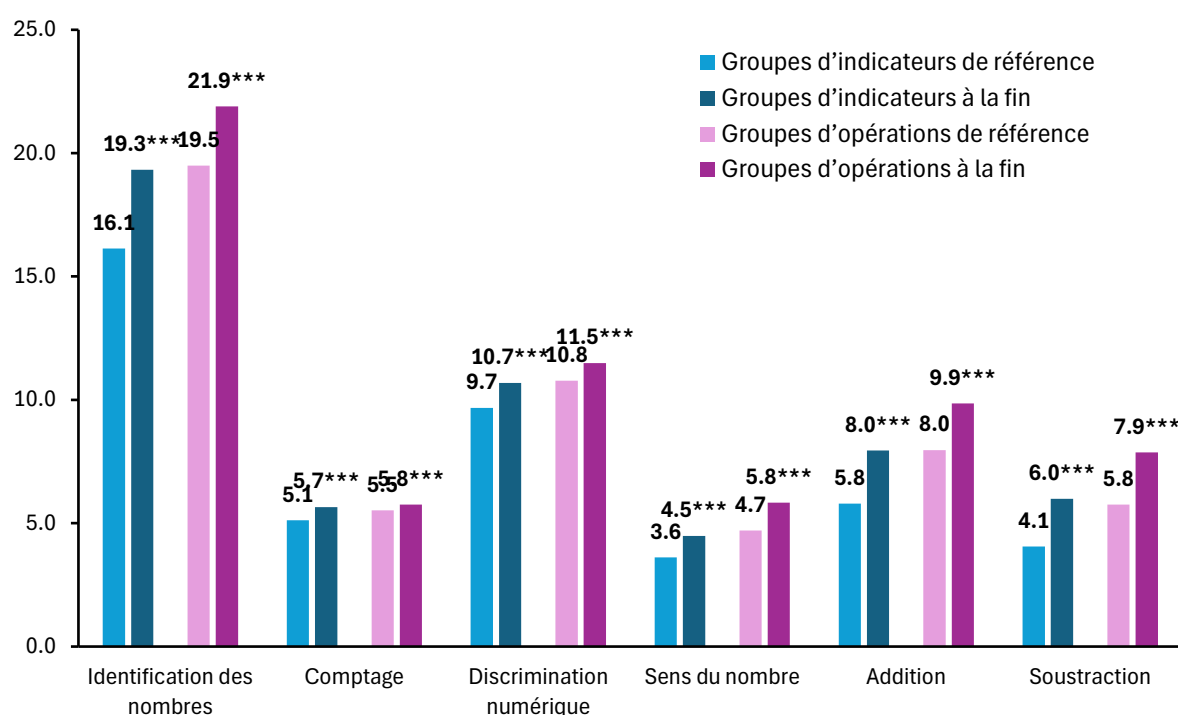
Remarque : Tous les effets du traitement sont normalisés. Les erreurs types sont indiquées entre parenthèses. Les contrôles comprennent le niveau scolaire, les variables indicatrices pour les filles, les élèves bénéficiant d'un soutien scolaire externe et la langue dans laquelle l'enquête a été menée. Le concept de décodage est l'effet moyen du traitement de trois sous-tests de décodage, notamment la connaissance des sons des lettres, le décodage de pseudo-mots et le décodage de mots familiers. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Le **tableau 15** présente les scores moyens obtenus à chaque sous-test de mathématiques pour les groupes « nombres » et « opérations ». En matière **d'identification** des nombres, les deux groupes ont enregistré des progrès significatifs, passant respectivement de 16,1 à 19,3 et de 19,7 à 22,0 sur 24 ($p < 0,01$). En matière de **comptage**, les deux groupes ont enregistré des progrès significatifs ($p < 0,01$), bien que ceux-ci aient été modestes. Le groupe « nombres

» est passé de 5,1 à 5,7 sur 6, tandis que le groupe « opérations » est passé de 5,5 à 5,8. Là encore, les élèves des deux groupes ont enregistré une amélioration légère mais significative dans le sous-test **de discrimination des chiffres**, le groupe « nombres » passant de 9,7 à 10,7 sur 12, tandis que le groupe « opérations » est passé de 10,9 à 11,5 ($p < 0,01$). En ce qui concerne **le sens des nombres**, le groupe « nombres » est passé de 3,6 au début de l'étude à 4,5 à la fin, sur une échelle de 12, tandis que le groupe « opérations » a vu son score passer de 4,8 à 5,9 ($p < 0,01$). Les deux groupes ont également enregistré des progrès significatifs dans les sous-tests **d'addition** et **de soustraction**, le groupe « nombres » passant respectivement de 5,8 à 8,0 et de 4,1 à 6,0 sur 12, tandis que le groupe « opérations » a amélioré ses scores en addition de 8,1 à 9,9 et en soustraction de 5,9 à 7,9 ($p < 0,01$).

Tableau 1515. Résultats en mathématiques du groupe de traitement par groupe mathématique



Remarque : les groupes mathématiques ont été déterminés par le test de sélection ARED (voir annexe A). $N = 141$ pour le groupe des nombres et 166 pour le groupe des opérations.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tests de robustesse

Nous avons effectué plusieurs tests de robustesse. Tout d'abord, pour tenir compte de la structure imbriquée des données (élèves au sein des écoles) et de la corrélation intra-scolaire, nous avons inclus un modèle à effets aléatoires. Les résultats du modèle à effets aléatoires sont similaires aux résultats principaux (**tableau D5, annexe D**). Ensuite, nous avons ajouté des variables de contrôle supplémentaires pour les variables observables au niveau des écoles et des élèves qui ont été recueillies au début de l'étude et qui présentaient des différences significatives entre les écoles et les élèves du groupe de traitement et ceux du groupe de comparaison : la présence d'un lavabo dans l'école et la disponibilité de savon dans ce lavabo, ainsi que le nombre moyen de jours de cours manqués au cours de la semaine précédente,

tel que déclaré par les élèves eux-mêmes. L'ajout de ces contrôles n'a pas eu d'incidence significative sur les résultats (**tableau D6, annexe D**). Ensuite, nous avons effectué les mêmes régressions en utilisant uniquement les élèves qui étaient présents au début et à la fin de l'étude (**tableau D7, annexe D**). Les régressions sur les échantillons appariés ont donné des résultats très similaires, à l'exception de la compréhension écrite (qui n'est plus statistiquement significative) et de l'identification des chiffres (marginale significative). Nous avons également effectué les régressions sur l'échantillon complet sans les élèves parlant le sérère (**tableau D8, annexe D**). Les régressions ont suivi un schéma similaire aux autres, tous les résultats de décodage étant significatifs et positifs. Cependant, dans ces régressions, la compréhension orale a eu un effet négatif et la conscience phonologique un effet positif, tandis qu'aucun des effets du traitement sur les mathématiques n'était significatif.

Effets hétérogènes du traitement

Dans la présente section, nous examinons les effets hétérogènes du traitement selon le genre, la langue et la région. Il convient de noter que la taille de notre échantillon ne permet de détecter que les effets globaux et que l'analyse statistique a une capacité limitée à détecter les effets hétérogènes.

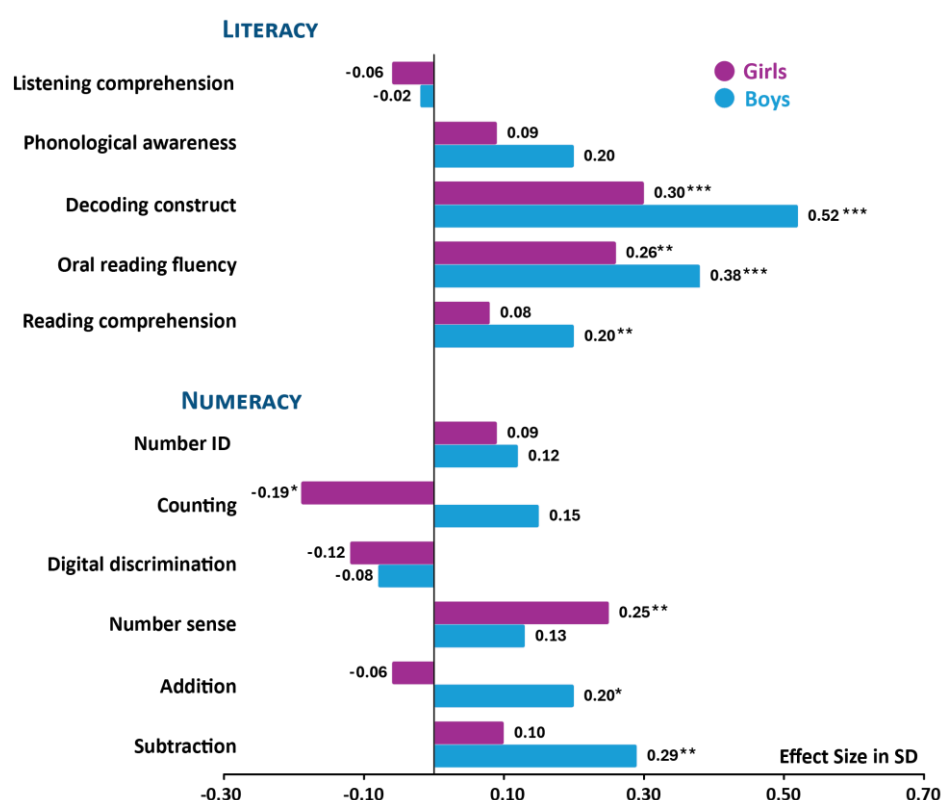
Tout d'abord, nous avons examiné les effets du traitement selon le genre. Comme le montre **le tableau 16**, *Ndaw Wune* a eu des effets positifs chez les filles comme chez les garçons, les garçons enregistrant toutefois des effets de traitement systématiquement plus élevés que les filles. Pour les principaux résultats en lecture-écriture mentionnés précédemment (compétences de décodage, notamment connaissance des sons des lettres, décodage de pseudo-mots et familiers, et fluidité de lecture à voix haute), garçons et filles présentent des effets significatifs. Au départ, les filles du groupe de traitement avaient des scores statistiquement beaucoup plus élevés que les garçons du groupe de traitement pour l'addition, la soustraction et deux des trois résultats de décodage. L'ampleur plus élevée des effets observés chez les garçons pourrait s'expliquer par une plus grande marge de progression, compte tenu de leurs niveaux initiaux, relativement à ceux des filles. En ce qui concerne les mathématiques, les effets variaient davantage selon le genre. Les filles ont eu un effet négatif sur l'exercice de comptage ($-0,19$ SD, $p < 0,10$), tandis que les garçons n'ont eu aucun impact significatif dans un sens ou dans l'autre. Pour le sous-test de sens du nombre, les filles enregistrent un effet positif ($0,25$ écart-type, $p < 0,05$), contrairement aux garçons. En revanche, les garçons ont eu des impacts positifs et significatifs tant pour l'addition ($0,2$ écart-type, $p < 0,05$) que pour la soustraction ($0,29$ écart-type, $p < 0,05$).

Ensuite, nous examinons les impacts selon la langue d'enseignement dans le **tableau 17**. Comme indiqué dans les résultats présentés pour la fluidité sémantique, la langue d'enseignement et la langue dominante d'un élève sont de bons indicateurs l'une de l'autre, étant donné le faible niveau de décalage entre elles. Nous présentons les résultats concernant l'impact pour les élèves wolophones et pulaar.²⁶ **Les données montrent que les élèves wolophones concentrent l'essentiel des gains associés au programme *Ndaw Wune*.** L'impact

²⁶ Nous n'avons pas accès aux informations relatives à la langue d'enseignement des écoles avant que l'échantillonnage ne soit effectué. Au final, nous disposons de données pour 12 écoles sérères du groupe expérimental et une seule école sérère du groupe témoin. Cet échantillon n'était pas suffisamment large pour permettre des comparaisons fiables entre les groupes expérimental et témoin. L'échantillon comprenait 12 écoles wolof du groupe expérimental et 23 écoles wolof du groupe témoin, ainsi que 11 écoles pulaar du groupe expérimental et du groupe témoin.

sur les élèves wolophones était significatif et positif pour la connaissance des sons des lettres (0,27 SD, $p < 0,01$), le décodage des pseudo-mots (0,5 SD, $p < 0,01$), le décodage de mots familiers (0,69 SD, $p < 0,01$) et la fluidité de lecture à voix haute (0,53 SD, $p < 0,01$), tandis que les élèves pulaarophones n'ont eu que des impacts significatifs sur la connaissance des sons des lettres (0,43 SD, $p < 0,10$). Cependant, les élèves wolophones ont également eu un effet négatif marginalement significatif sur la compréhension orale (-0,3 SD, $p < 0,10$).²⁷ Les élèves wolophones n'ont pas eu d'impact significatif sur les exercices de mathématiques, et les locuteurs pulaar ont obtenu des résultats négatifs dans un sous-test, celui du comptage (-0,37 SD, $p < 0,10$).²⁸

Tableau 1616 . Effets du traitement selon le genre



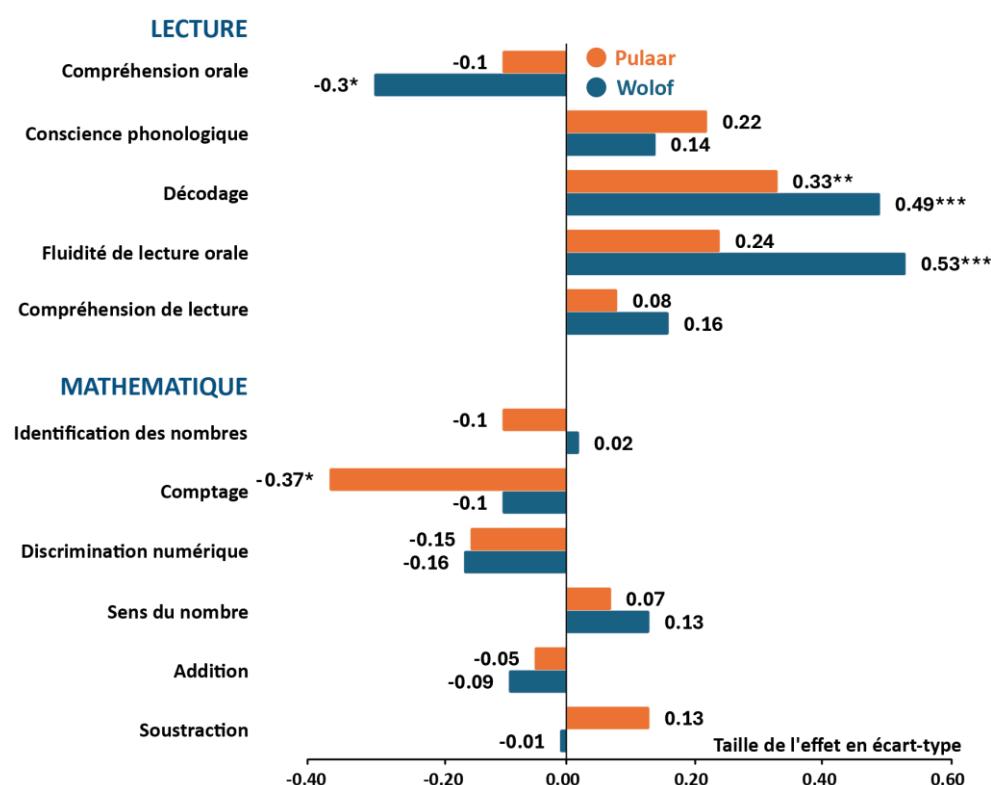
Remarque : Tous les effets du traitement sont standardisés. Les contrôles comprennent le niveau scolaire, le fait que l'élève ait bénéficié ou non d'un soutien scolaire externe et la langue dans laquelle l'enquête a été menée.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

²⁷ Au départ, les élèves des écoles de traitement wolof avaient des scores élevés au test de compréhension orale (4,4 sur 6 contre 3,4 sur 6 pour les élèves de comparaison). Bien que les deux groupes aient progressé entre le début et la fin de l'étude, un effet plafond a peut-être empêché les élèves du groupe de traitement de progresser à un rythme identique ou supérieur à celui des élèves du groupe de comparaison, ce qui a conduit à une estimation négative de l'impact.

²⁸ Nous avons vérifié l'échantillon apparié et les résultats sont similaires à ceux de l'échantillon complet. De plus, il y avait très peu de décalage entre la langue dominante d'enseignement et la langue dominante parlée par l'élève (mesurée par la fluidité sémantique). Une explication possible des impacts négatifs sur certaines sous-compétences mathématiques pourrait être que, comme le montrent les commentaires des enseignants, les exercices de calcul étaient difficiles et parfois difficiles à comprendre pour les élèves. De plus, étant donné que la discrimination numérique semble être la sous-compétence qui a l'impact négatif le plus constant, il serait prudent de revoir des exercices simples tels que le comptage et la comparaison des grandeurs numériques afin de renforcer et de consolider ces concepts.

Tableau 1717 . Effets du traitement selon la langue d'enseignement



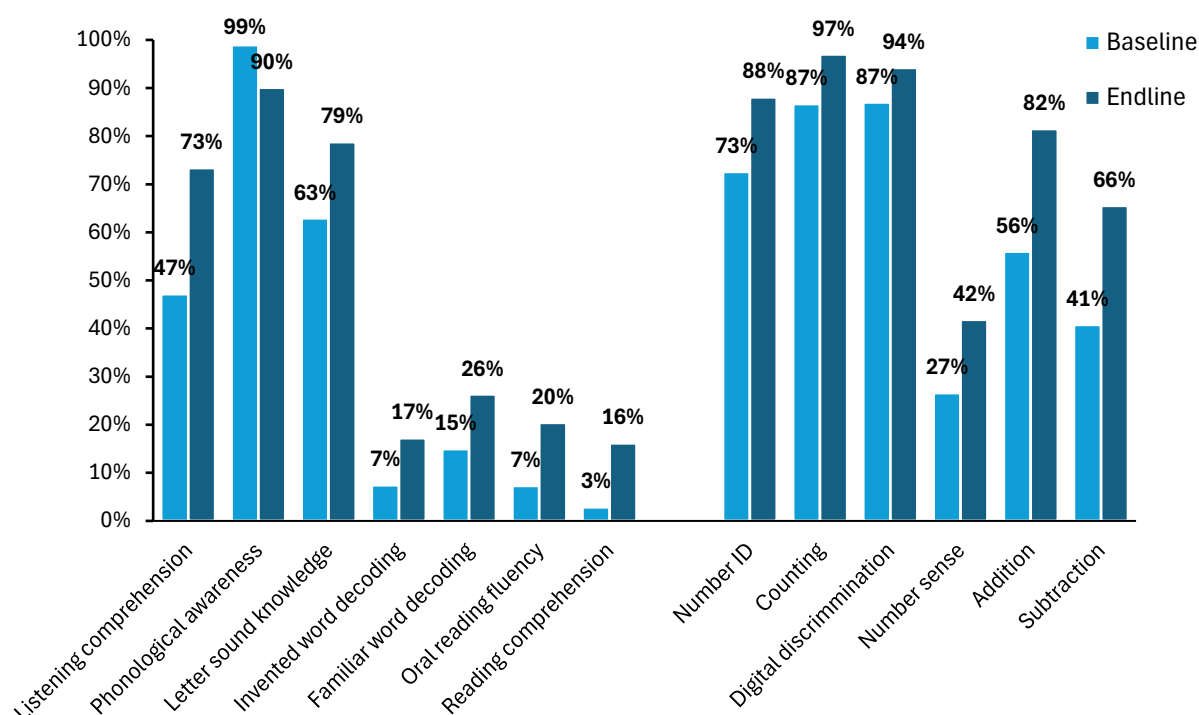
Remarque : tous les effets du traitement sont normalisés.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Comme nous ne disposons pas de données comparatives suffisantes pour établir des comparaisons fiables entre le groupe de traitement et le groupe de comparaison pour les élèves sérérophones, nous avons procédé à une analyse pré/post limitée aux élèves du groupe de traitement uniquement, en comparant les scores de référence et finaux des élèves du groupe de traitement pour chaque sous-test. **Le tableau 18** présente les scores moyens EGRA et EGMA de référence et finaux exprimés en pourcentage. Dans la quasi-totalité des sous-tests, les élèves ont enregistré une amélioration positive et significative entre les scores de référence et finaux.²⁹

²⁹ Pour certains sous-tests, il y a eu des effets plafonds possibles qui ont conduit à des gains plus modestes, mais néanmoins significatifs, tels que le comptage et la discrimination des chiffres, qui avaient des scores de référence élevés. Pour un sous-test, la conscience logique, les élèves sérères ont obtenu un score moyen de 99 % au départ et ont en fait enregistré une baisse significative à la fin, même si leur score moyen de 90 % restait assez élevé.

Annexe 1818 . Moyennes initiales et finales pour le groupe de traitement des élèves sérères



Remarque : N = 192 au début et 192 à la fin (186 pour la compréhension écrite).

Nous examinons ensuite les impacts par région.³⁰ Le **tableau 19** présente les résultats des impacts par département. Les impacts les plus marqués de *Ndaw Wune* ont été observés à Kaolack, où les élèves ont bénéficié d'impacts positifs dans quatre des six sous-tests de lecture-écriture.³¹ L'effet du traitement a été particulièrement fort pour le décodage de mots familiers (0,59 SD, $p < 0,01$), suivi de la fluidité de lecture à voix haute (0,48 SD, $p < 0,05$) et du décodage des pseudo-mots (0,45 SD, $p < 0,01$). Bien que les élèves de Saint-Louis et de Matam aient montré une amélioration dans la plupart des sous-tests entre le début et la fin de l'étude, les effets du traitement ont été nettement plus modestes à Saint-Louis et largement absents à Matam.³² Étant donné que l'ensemble des élèves pulaarophones de l'échantillon fréquentent des écoles dans ces deux régions, il n'est pas surprenant que leurs performances n'aient pas été à la hauteur des résultats plus probants observés dans les autres départements.

³⁰ Nous ne disposions pas d'un échantillon suffisamment important à Diourbel pour établir des comparaisons fiables entre le groupe traité et le groupe témoin.

³¹ La compréhension orale a eu des effets négatifs et significatifs à Kaolack. Nous avons vérifié l'échantillon apparié à Kaolack et il présentait un effet du traitement de -0,25, significatif à 5 %, ce qui correspondait au résultat obtenu pour l'échantillon complet. Comme nous n'avions pas la puissance statistique nécessaire pour les effets des sous-groupes, il est possible que ce soit un artefact dû au fait que ce sous-échantillon n'est pas suffisamment puissant.

³² Nous avons vérifié l'inadéquation linguistique à Matam et avons constaté une très faible inadéquation entre la langue d'enseignement et la langue parlée par l'élève. Il convient de noter que l'analyse n'était pas suffisamment puissante pour détecter les effets au niveau des sous-groupes. Ainsi, l'absence d'impacts à Matam pourrait être due à la puissance insuffisante de l'étude. Enfin, en l'absence de données qualitatives approfondies sur la fidélité de la mise en œuvre, nous ne sommes pas en mesure de comprendre les raisons qui expliquent les impacts faibles ou inexistants à Matam.

Figure 191920 . Effets du traitement selon la région

Résultat	Kaolack		Matam		Saint-Louis	
	Effet du traitement	N	Effet du traitement	N	Effet du traitement	N
Compréhension orale (sur 6)	-0,2** (0,09)	630	0,01 (0,15)	511	-0,18 (0,15)	542
Conscience phonologique (sur 10)	0,25 (0,09)	630	0,29 (0,2)	511	0,17 (0,16)	542
Décodage conceptuel	0,41*** (0,11)	630	0,20 (0,14)	511	0,40 (0,22)	542
• Connaissance des sons des lettres (sur 100)	0,19*** (0,08)	63	0,43 (0,29)	511	0,32** (0,14)	542
• Décodage de pseudo-mots (sur 50)	0,45 (0,14)	630	0,13 (0,12)	511	0,32 (0,25)	542
• Décodage de mots familiers (sur 50)	0,59*** (0,18)	630	0,06 (0,15)	511	0,57 (0,31)	542
Fluidité de lecture à voix haute (sur 50)	0,48 (0,2)	630	0,03 (0,16)	511	0,31 (0,30)	542
Compréhension écrite (sur 6)	0,09 (0,2)	619	0,03 (0,13)	508	0,09 (0,18)	529
Numéro d'identification (sur 24)	0,13 (0,07)	630	-0,09 (0,19)	511	-0,01 (0,10)	542
Comptage (sur 6)	0,17 (0,11)	630	-0,45*** (0,13)	511	-0,09 (0,16)	542
Discrimination des chiffres (sur 12)	0,08 (0,06)	630	-0,13 (0,14)	511	-0,15 (0,14)	542
Sens des nombres (sur 12)	0,22 (0,14)	630	-0,16 (0,23)	511	0,22 (0,15)	542
Addition (sur 12)	0,13 (0,12)	630	-0,01 (0,12)	511	-0,14 (0,10)	542
Soustraction (sur 12)	0,24 (0,14)	630	0,08 (0,11)	511	0,04 (0,13)	542

Remarque : tous les effets du traitement sont normalisés. Les erreurs types sont indiquées entre parenthèses. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

6.2. Analyse des données de suivi

Comme décrit à la section 5.2, les chercheurs d'AIR ont également analysé les commentaires des enseignants recueillis par ARED dans le cadre du suivi du programme, pour compléter les

résultats quantitatifs. Il s'agit d'informations anecdotiques fournies par les enseignants eux-mêmes, qui n'ont été mentionnées que par moins de 1 à 7 % d'entre eux dans chaque cas. Les principales difficultés récurrentes qui ressortent de ces données sont les suivantes : un effectif d'élèves perçu comme élevé rendant la gestion de la classe plus difficile ; des taux d'absences élevés ; l'indisponibilité ou l'absence de certains matériels nécessaires à l'enseignement ; des inexactitudes dans les supports ; et des supports de mathématiques jugés difficiles à utiliser. Compte tenu du faible nombre d'occurrences, nous les présentons ici de manière succincte, à titre d'éléments contextuels venant compléter les résultats d'impact et à prendre en considération dans la perspective d'une mise à l'échelle du programme *Ndaw Wune*.

6.3. Analyse des coûts

Sur la base de la méthodologie d'analyse des coûts décrite à la section 4.3, nous avons estimé les coûts de mise en œuvre de *Ndaw Wune Année Scolaire*, que nous présentons ci-dessous, par question d'évaluation. Nous présentons tous les coûts en francs CFA (F CFA) et en dollars américains (USD), en utilisant un taux de conversion de 580 F CFA pour 1 USD.

Ndaw Wune Année Scolaire

Nous avons estimé le coût global du programme *Année Scolaire* 2024-2025 à 251 141 271 F CFA (433 002 USD) et le coût direct de sa mise en œuvre à 226 027 144 F CFA (75 342 USD).³³ Ces coûts couvrent 65 écoles réparties dans six IEF, soit 100 classes et 3 000 élèves au total. Le coût global par élève est de 83 714 F CFA (144 USD) et le coût direct par élève est de 389 702 F CFA (130 USD).

Tableau 21. Coûts totaux de l'année scolaire

Coût global	Coût (F CFA)	Coût (USD)
Coût total	251 141 271	433 002
Coût par étudiant	83 714	144

Notre estimation des coûts est nettement inférieure au budget annuel d'ARED pour *l'année scolaire* (544 843 dollars américains), ce qui est courant lorsque l'on compare les dépenses réelles aux dépenses budgétisées. Cette différence peut s'expliquer par de légers ajustements dans la mise en œuvre effective du programme, ainsi que par les hypothèses retenues dans l'analyse des coûts. Par exemple, l'analyse des coûts utilise le protocole G50 actuel pour calculer les frais de déplacement, qui prévoit des taux de remboursement inférieurs à ceux du protocole G50 qui était en vigueur pendant la majeure partie de l'année scolaire 2024-2025.

Nous examinons ensuite les coûts par composante du programme. Pour ce faire, nous avons divisé le programme *Année scolaire* 2024-2025 en quatre composantes :

1. Formation initiale des enseignants
2. Enseignement direct

³³ Les coûts directs de mise en œuvre correspondent aux coûts globaux moins les frais généraux de 10 %.

3. Journées pédagogiques

4. Gestion, planification et évaluation

Formation initiale des enseignants. Avant le démarrage du programme chaque année, ARED organise une formation à l'intention de l'ensemble des enseignants chargés d'assurer l'enseignement de *Ndaw Wune* en classe. Cette formation consiste en une « journée d'harmonisation » pour tous les formateurs d'enseignants dans un lieu central (Diourbel pour l'année scolaire 2024-2025), suivie d'une journée de déplacement, puis de quatre jours de formation dans chaque IEF. Les coûts comprennent le temps de travail du personnel d'ARED, des consultants et des inspecteurs, les frais de transport et les indemnités journalières pour les repas et l'hébergement, le coût annualisé des projecteurs et des ordinateurs portables, ainsi que les frais généraux estimés à 10 % par ARED.

Enseignement direct. Dans la catégorie « enseignement direct », nous incluons les coûts liés à la mise en œuvre du programme *Ndaw Wune* en classe et au suivi continu du programme. Les pratiques de suivi d'ARED pendant l'année scolaire 2024-2025 comprenaient des rapports réguliers des superviseurs et des rapports ultérieurs du personnel d'ARED. Les coûts comprennent les frais de personnel d'ARED, les allocations versées aux enseignants, aux superviseurs et aux inspecteurs, ainsi que les frais d'impression et de livraison de tout le matériel pédagogique. Les coûts liés aux installations ne sont pas inclus, car ils ne constituent pas un coût pour le programme, et les frais généraux sont estimés à 10 %.

Journées pédagogiques. Deux fois par année scolaire, ARED organise des « journées pédagogiques », ou formations de remise à niveau, pour les enseignants participant au programme. Dans le cadre de ces journées pédagogiques, le personnel d'ARED se rend dans chacune des IEF pour observer les classes et rencontrer les directeurs d'école. Ensuite, il organise des formations de remise à niveau auxquelles participent les enseignants, les superviseurs, les points focaux et les directeurs d'école. Au cours de l'année scolaire 2024-2025, ARED a organisé les journées pédagogiques en trois groupes (un comprenant les IEF de Kaolack et Bambey, un comprenant les IEF de Saint-Louis et Podor, et un comprenant les IEF de Kanel et Matam). Nos estimations de coûts pour les journées pédagogiques comprennent le temps de travail du personnel d'ARED, des animateurs et des inspecteurs, les frais de transport et les indemnités journalières pour les repas et l'hébergement, ainsi que les frais généraux estimés à 10 %.

Gestion, planification et évaluation. Cette catégorie regroupe l'ensemble des activités résiduelles du programme mises en œuvre durant les **trois mois** de l'année scolaire au cours desquels *Année Scolaire* n'est pas en session. Elle comprend les tâches administratives liées à la gestion du programme et à la planification des activités, ainsi que les coûts annuels d'évaluation. Les coûts d'évaluation incluent une journée annuelle de formation des enquêteurs dans un site central, ainsi que deux jours de collecte de données dans chaque IEF, par vague. Au sein de cette catégorie, sont comptabilisés : les coûts de personnel du staff d'ARED ; une indemnité de 30 000 FCFA par enquêteur et par jour de collecte ; les frais de transport et les indemnités journalières (repas et hébergement) ; les coûts annualisés des tablettes et de l'abonnement au logiciel Survêtit ; les frais de location de salle pour la formation des enquêteurs ; ainsi que des frais généraux estimés à 10 %.

Le tableau 21 présente les coûts de chaque composante du programme et le pourcentage des coûts totaux du programme alloués à chaque composante. L'enseignement direct

représente la plus grande partie des coûts totaux du programme (~62 %), suivi par la gestion, la planification et l'évaluation (~23 %). Les deux cycles de journées pédagogiques représentent environ 7 % et la semaine de formation des enseignants environ 8 % des coûts totaux.

Tableau 22. Coûts de l'Année scolaire par composante du programme

Composante	Coût (F CFA)	Coût (USD)	% du total
Formation initiale des enseignants	19 867 653	34 255	7,91
Enseignement direct	155 449 958	268 017	61,90
Journées pédagogiques	18 713 517	32 265	7,45
Gestion, planification et évaluation	57 110 144	98 466	22,74

Nous analysons les coûts totaux par catégorie de ressources dans le tableau 22. La plus grande partie des ressources est consacrée aux frais de personnel, les salaires du personnel d'ARED représentant la composante majoritaire de cette catégorie. Parmi les coûts non liés au personnel, les frais de déplacement constituent la part prédominante. Les frais généraux sont estimés à 10 %, sur la base des échanges avec les équipes d'ARED.

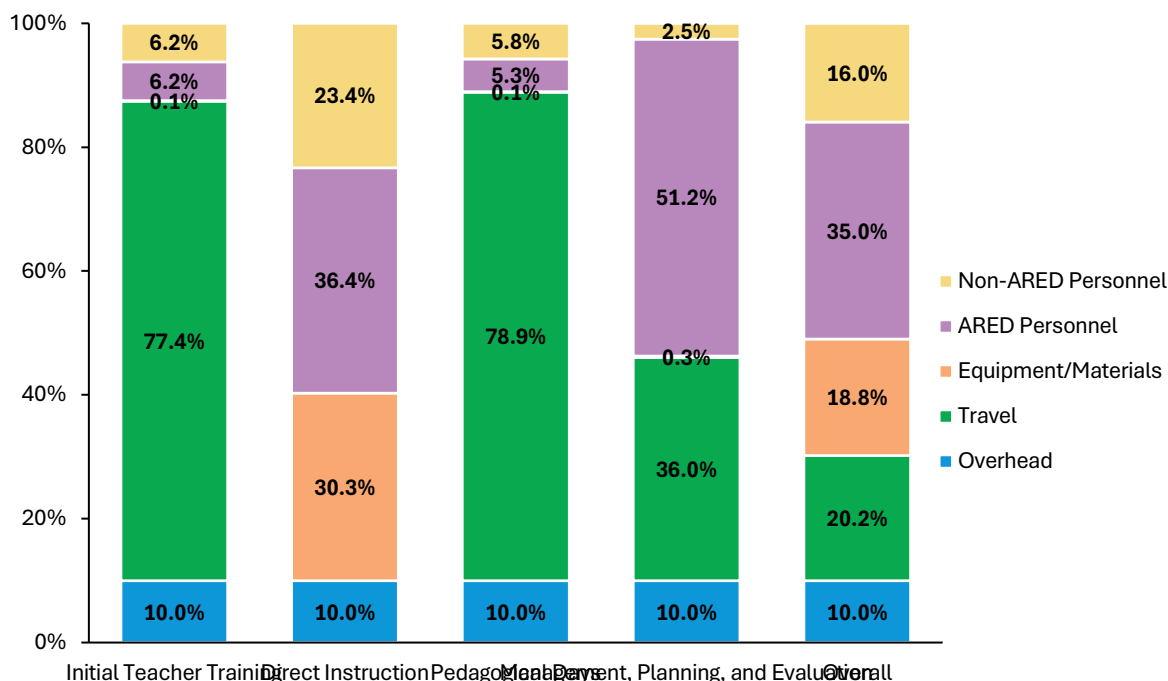
Tableau 23. Coûts d'Année Scolaire par catégorie de ressources

Catégorie de ressources	Coût (F CFA)	Coût (USD)	% du total
Personnel	128 065 591	220 803	50,99
Personnel d'ARED	88 005 591	151 734	35,04
Consultants/facilitateurs	1 800 000	3 103	0,72
Enseignants	30 000 000	51 724	11,95
Superviseurs	6 000 000	10 345	2,39
Inspecteurs	820 000	1 414	0,33
Évaluateurs	1 440 000	2 483	0,33
Dépenses hors personnel	97 961 553	168 899	39,01
Déplacements	50 705 099	87 423	20,19
Équipement/Matériel	47 231 454	81 434	18,81
Installations	25 000	43	0,01
Frais généraux	25 114 127	43 300	10,00
Total	251 141 271	433 002	100,00

Dans le tableau 23, nous analysons la répartition des ressources au sein de chaque composante du programme, en la comparant à la structure globale des coûts. Les frais de déplacement constituent la part la plus importante des coûts pour la formation initiale des enseignants et les journées pédagogiques, tandis que les coûts de personnel représentent la

majorité des dépenses pour les composantes enseignement direct et gestion, planification et évaluation. Un tableau complet des coûts par composante du programme et par catégorie de ressources figure à l'annexe C.

24 Tableau 25. Répartition des coûts pour l'Année scolaire par composante du programme



Dans la mesure où ARED applique les mêmes barèmes d'indemnités aux enseignants, superviseurs et inspecteurs dans l'ensemble des régions, et où le protocole G50 fixe des taux de remboursement des déplacements et des indemnités pour le personnel public homogènes à l'échelle nationale, nous n'observons pas de variations substantielles des coûts selon la localisation du programme. Les principales différences concernent les **frais de transport**, qui augmentent pour le personnel d'ARED en fonction de l'éloignement des sites de mise en œuvre par rapport à Dakar, ainsi que, dans certains cas, les **coûts d'impression**. ARED a indiqué que certains supports (par exemple, les guides enseignants et les cahiers de mathématiques des élèves) présentent des coûts d'impression identiques quelle que soit la langue, tandis que d'autres (notamment les cahiers de lecture des élèves) varient selon la langue, en raison de différences dans le nombre de pages.

Scénarios de mise à l'échelle

Enfin, nous présentons des projections de coûts correspondant à quatre scénarios de mise à l'échelle de *Ndaw Wune Année Scolaire*. Dans chacun de ces scénarios, le programme est étendu à 50 000 élèves, répartis dans 1 670 classes, au sein des mêmes six IEF que dans le scénario de référence (Podor, Matam, Saint-Louis, Kaolack, Bambey et Kanel). Dans le cadre des scénarios de mise à l'échelle, ARED ajoute deux employés salariés, les coûts des matériels à l'unité sont réduits conformément aux bases de données de la Banque mondiale pour l'impression à grande échelle, tous les coûts d'évaluation sont supprimés, ainsi que tous les coûts de personnel liés à l'élaboration des supports pédagogiques.

- Scénario 1 : Mise à l'échelle à 50 000 élèves dans les six IEF, sans modification du volume horaire d'enseignement, de la formation initiale des enseignants ni des journées pédagogiques. Les directeurs d'école assument les fonctions de supervision.
- Scénario 2 : Mise à l'échelle à 50 000 élèves dans les six IEF, sans modification du volume horaire d'enseignement. La formation initiale des enseignants est réduite à quatre demi-journées. Les indemnités des superviseurs et des inspecteurs sont supprimées.
- Scénario 3 : Mise à l'échelle à 50 000 élèves dans les six IEF, avec l'ensemble des enseignements dispensés durant les heures officielles de classe, à raison d'un maximum d'une heure de mathématiques et une heure de lecture-écriture par semaine. La formation initiale des enseignants est réduite à quatre demi-journées. Les indemnités des superviseurs, des inspecteurs et des enseignants sont supprimées.
- Scénario 4 : Reprise de l'ensemble des hypothèses du Scénario 3, avec, en outre, l'utilisation de supports pédagogiques existants, entraînant la suppression de tous les coûts liés à l'impression et à la distribution des supports pédagogiques.

Scénario 1. Dans le premier scénario, les indemnités des enseignants sont réduites à 25 000 F CFA par mois. Le rôle de superviseur est assumé par les directeurs d'école, à raison de 65 superviseurs pour 100 écoles, pour une indemnité de 30 000 F CFA par trimestre. Au total, 112 animateurs sont nécessaires pour les formations des enseignants et les journées pédagogiques pour former 1 670 enseignants dans 56 districts, et ils sont rémunérés 10 000 F CFA par jour. Comme les formations ont lieu dans chaque district, les frais de déplacement des non-résidents sont remplacés par les frais de déplacement des résidents. On suppose qu'un ordinateur portable et un projecteur sont nécessaires pour chaque formation (c'est-à-dire un par district). Comme le montre le **tableau 24**, le coût estimé pour le scénario 1 est de 767 650 752 F CFA (1 323 536 USD). Le coût total par élève est de 15 353 F CFA (26 USD) et les coûts directs par élève sont de 13 818 F CFA (24 USD).

Tableau 26. Coûts estimés pour le scénario 1

Catégorie de ressources	Coût (F CFA)	Coût (USD)
Personnel	377 768 617	651 325
Personnel ARED	87 148 617	150 256
Facilitateurs	6 720 000	11 586
Enseignants	250 500 000	431 897
Superviseurs	32 580 000	56 172
Inspecteurs	820 000	1 414
Dépenses hors personnel	313 117 060	539 857
Déplacements	225 686 026	389 114

Équipement/Matériel	87 431 035	150 743
Frais généraux	76 765 075	132 354
Total	767 650 752	1 323 536
Par étudiant	15 353	26

Scénario 2. Dans le deuxième scénario, nous conservons la plupart des hypothèses du scénario 1, mais réduisons la formation initiale des enseignants à quatre demi-journées au lieu de quatre journées complètes. Nous supprimons également les indemnités versées aux inspecteurs et aux superviseurs. Toutefois, les inspecteurs/points focaux continuent de percevoir une indemnité journalière de 10 000 F CFA pour assister aux formations. Comme le montre le **tableau 25**, le coût estimé du scénario 2 est de 638 388 975 F CFA (1 100 671 USD). Le coût total par élève est de 12 768 F CFA (22 USD) et les coûts directs par élève sont de 11 491 F CFA (20 USD).

Tableau 2728. Coûts estimés pour le scénario 2

Catégorie de ressources	Coût (F CFA)	Coût (USD)
Personnel	344 888 617	594 636
Personnel ARED	87 148 617	150 256
Facilitateurs	6 720 000	11 586
Enseignants	250 500 000	431 897
Superviseurs	0	0
Inspecteurs	520 000	897
Dépenses hors personnel	229 661 460	395 968
Déplacements	142 230 426	245 225
Équipement/Matériel	87 431 035	150 743
Frais généraux	63 838 897	110 067
Total	638 388 975	1 100 671
Par étudiant	12 768	22

Scénario 3. Dans ce scénario, nous supposons que tous les cours ont lieu pendant les heures normales de classe ; nous supprimons l'indemnité versée aux enseignants et réduisons le nombre d'heures d'enseignement de six à deux au maximum. Comme le montre le **tableau 26**, le coût estimé pour le scénario 3 est de 360 055 641 F CFA (620 786 USD). Le coût total par élève est de 7 201 F CFA (12 USD) et les coûts directs par élève sont de 6 481 F CFA (11 USD).

Tableau 29. Coûts estimés pour le scénario 3

Catégorie de ressources	Coût (F CFA)	Coût (USD)
-------------------------	--------------	------------

Personnel	94 388 617	162 739
Personnel ARED	87 148 617	150 256
Facilitateurs	6 720 000	11 586
Enseignants	0	0
Superviseurs	0	0
Inspecteurs	520 000	897
Dépenses hors personnel	229 661 460	395 968
Déplacements	142 230 426	245 225
Équipement/Matériel	87 431 035	150 743
Frais généraux	36 005 564	62 079
Total	360 055 641	620 786
Par étudiant	7 201	12

Scénario 4. Dans ce scénario, les élèves utilisent les supports existants, ce qui évite les coûts d'impression/de livraison des supports pédagogiques. **Le tableau 27** montre que la suppression des coûts liés au matériel scolaire réduit encore les coûts globaux à 263 343 808 F CFA (454 041 USD) et les coûts par élève à 5 267 F CFA (9 USD). Les coûts de mise en œuvre directe sont réduits à 237 009 427 F CFA (4 008 637 USD) et à 4 740 F CFA (8 USD) par élève.

Tableau 30. Coûts estimés pour le scénario 4

Catégorie de ressources	Coût (F CFA)	Coût (USD)
Personnel	94 388 617	162 739
Personnel ARED	87 148 617	150 256
Facilitateurs	6 720 000	11 586
Enseignants	0	0
Superviseurs	0	0
Inspecteurs	520 000	897
Non liés au personnel	142 620 810	245 898
Déplacements	142 230 426	245 225
Équipement/Matériel	390 385	673
Frais généraux	26 334 381	45 404
Total	263 343 808	454 041
Par étudiant	5 267	9

6.4. Analyse coût-efficacité

En croisant les résultats de l'analyse des coûts de *Ndaw Wune Année Scolaire* avec ceux de l'évaluation d'impact, nous présentons des estimations de coût-efficacité selon une approche d'inventaire d'impact. Dans l'ensemble, nous constatons que pour un coût de 83 714 F CFA (soit 144 USD) par élève, *Ndaw Wune Année Scolaire* a généré des effets positifs sur plusieurs acquis fondamentaux en lecture-écriture et en mathématiques, comme l'illustre le **tableau 28**.

Tableau 31. Inventaire illustratif de l'impact

PROGRAMME	INVESTISSEMENT	IMPACTS	
FORMATION DES ENSEIGNANTS	TOTAL 251,141,271 FCFA (433,002 USD)	LECTURE	MATHÉMATIQUES
INSTRUCTION DIRECTE		 	 
JOURNÉES PÉDAGOGIQUES	PAR ÉTUDIANT 83,714 FCFA (144 USD)	 	 
PLANIFICATION, SUIVI ET ÉVALUATION		 	

Dans le **tableau 29**, nous avons évalué de manière linéaire l'ampleur des effets pour chaque résultat, pour refléter l'impact par tranche de 100 dollars US investis par élève dans *Ndaw Wune Année Scolaire*. Nous indiquons l'impact par tranche de 100 dollars pour faciliter les comparaisons de coût-efficacité entre plusieurs programmes d'enseignement de rattrapage. Comme il est d'usage dans ce type d'analyse, il convient de préciser que ce ratio n'implique pas une relation strictement linéaire entre le niveau d'investissement et l'ampleur des effets. Un investissement de 58 000 FCFA (100 USD) par élève est associé à des gains d'ampleur modérée pour la connaissance des sons des lettres, le décodage de mots et la fluidité de lecture à voix haute, ainsi qu'à des gains de plus faible ampleur pour la compréhension en lecture et les résultats en mathématiques.

Tableau 32. Impacts obtenus pour 100 dollars investis par élève

Compétence	Résultat (écart-type et IC à 90 %)
Connaissance des lettres et décodage des mots	0,29 SD (0,20–0,37)
Fluidité de lecture à voix haute	0,22 écart-type (0,09–0,35)
Compréhension écrite	0,10 SD (0,01–0,19)
Sens des nombres	0,13 SD (0,03–0,23)
Soustraction	0,14 SD (0,03–0,24)

7. Conclusions

Sur la base de notre évaluation, nous concluons que *Ndaw Wune Année Scolaire* constitue, dans l'ensemble, un programme efficace pour améliorer les acquis fondamentaux en **lecture-écriture** dans les langues nationales chez les élèves en difficulté des classes de CE1–CE2. Plus précisément, *Ndaw Wune* a généré des effets positifs et significatifs sur les sous-compétences de reconnaissance des mots ainsi que sur les compétences de compréhension en lecture dans les langues nationales. En revanche, aucun effet significatif n'a été observé sur la conscience phonologique ni sur les compétences de langage oral. Par ailleurs, *Ndaw Wune* a produit un effet plus modeste, mais néanmoins positif, sur les acquis fondamentaux des élèves en mathématiques. Nous examinons chacun de ces points ci-dessous dans le cadre théorique de l'apprentissage de la lecture dans le contexte multilingue du Sénégal.

Pour rappel, la théorie CFRA (présentée dans la section 2) postule que les acquis fondamentaux en lecture-écriture dépendent à la fois des sous-compétences de reconnaissance des mots et des sous-compétences de compréhension du langage, l'automatisation de ces composantes conduisant à la compréhension en lecture. Dans notre étude, les élèves du groupe de traitement ont obtenu des résultats positifs et significatifs dans plusieurs sous-compétences de reconnaissance des mots, notamment la connaissance des sons des lettres (traitement grapho-phonologique) (0,45 écart-type), les compétences de décodage (0,42 écart-type), la fluidité de lecture à voix haute (0,32 écart-type) et la compréhension en lecture (0,14 écart-type). Ces résultats indiquent que le programme *Ndaw Wune* est particulièrement efficace pour renforcer les composantes de reconnaissance des mots nécessaires au développement de la compréhension en lecture. La conscience phonologique augmente en amplitude, sans toutefois atteindre le seuil de significativité statistique, et aucun effet significatif n'est observé sur la compréhension orale. Compte tenu du contexte — dans lequel moins d'un tiers des élèves a accès à des opportunités d'apprentissage et de lecture dans une langue qu'ils comprennent et parlent bien, principalement le français (Nakamura & Holla, 2024) —, ces résultats constituent un indicateur fort : lorsque l'enseignement est systématique et explicite sur les sous-compétences de décodage et de reconnaissance des mots, la compréhension en lecture peut progresser dès lors que les enfants sont instruits dans des langues qu'ils utilisent et comprennent. Il est essentiel de noter que le succès de *Ndaw Wune* dans le développement de la compréhension écrite est probablement dû à la combinaison d'une pédagogie structurée efficace ciblant la reconnaissance des mots (c'est-à-dire un enseignement explicite de la phonétique, une portée et un ordre clairs des supports, des instructions fondées sur l'évaluation, ainsi qu'une disponibilité et des liens efficaces entre les manuels scolaires et les guides pédagogiques) et d'un enseignement dispensé aux enfants dans les langues qu'ils utilisent et comprennent. Ces résultats soulignent que le programme *Ndaw Wune* **favorise effectivement le développement des compétences fondamentales en lecture-écriture chez les élèves sénégalais de CE1–CE2.**

Deux des six sous-tests de **mathématiques** présentent des effets positifs et statistiquement significatifs, notamment le sens des nombres (0,19 écart-type) et la soustraction (0,20 écart-type). Les autres sous-tests ne montrent pas d'effets significatifs. Les gains observés en sens du nombre et en soustraction sont particulièrement importants, dans la mesure où ces compétences constituent le socle des apprentissages mathématiques aux premières années

du primaire. L'amélioration du sens des nombres indique que les élèves sont plus aptes à reconnaître des schémas, à comprendre les relations numériques et à résoudre des problèmes impliquant des nombres manquants, compétences essentielles pour tout apprentissage futur des mathématiques. L'impact positif sur la soustraction démontre les progrès réalisés par les élèves dans leur capacité à effectuer des opérations de base, ce qui constitue une étape cruciale vers la maîtrise d'opérations arithmétiques plus complexes. Cependant, quatre sous-tests de mathématiques, à savoir le comptage, la discrimination des chiffres, l'addition et l'identification des nombres, ne présentent pas d'effets statistiquement significatifs. Cela suggère que l'intervention est **efficace pour renforcer le raisonnement mathématique fondamental et les opérations de base, mais qu'elle pourrait nécessiter d'être affinée pour couvrir l'ensemble des compétences mathématiques des premières années**. Ces résultats mitigés soulignent l'importance d'adapter les stratégies et le matériel pédagogiques pour assurer un développement complet dans tous les domaines clés des mathématiques.

Impacts sur les sous-groupes. Comme indiqué précédemment, l'analyse des sous-groupes n'était pas suffisamment puissante pour détecter des effets significatifs. Nous présentons ici des conclusions suggestives basées sur les données.

1. Nous constatons que les résultats des garçons et des filles ont été influencés positivement par *Ndaw Wune*. Les garçons ont obtenu des résultats positifs et significatifs dans trois sous-tests de lecture-écriture, contre deux pour les filles, la différence portant sur la compréhension écrite. Les résultats des garçons étaient également systématiquement plus élevés. Les résultats sont plus mitigés en mathématiques, mais les garçons ont obtenu des effets positifs dans les sections addition et soustraction, tandis que les filles ont obtenu un effet positif dans le sens des nombres et un effet négatif dans le sous-test de comptage. Des études comparables montrent des résultats variables selon les contextes : par exemple, en Afrique du Sud, une évaluation a constaté un avantage persistant des filles (Fleisch et al., 2017). L'évaluation d'un programme de rattrapage au Niger n'a révélé aucune différence entre les sexes dans les effets du traitement (Maruyama & Kurosaki, 2021).
2. Nous n'avons pas accès aux informations concernant la langue d'enseignement dans les écoles avant que l'échantillonnage des écoles de comparaison ne soit effectué. Au final, nous disposons de données pour 12 écoles sérères du groupe de traitement et une seule école sérère du groupe de comparaison. Cet échantillon n'était pas suffisamment large pour permettre des comparaisons fiables entre les groupes expérimentaux et de comparaison. L'échantillon comprenait 12 écoles wolof du groupe de traitement et 23 écoles wolof du groupe de comparaison, ainsi que 11 écoles pulaar du groupe de traitement et du groupe de comparaison. Les élèves wolophones ont enregistré les gains les plus probants. Les données indiquent que les élèves wolophones ont généralement obtenu de meilleurs résultats aux sous-tests de lecture-écriture que les locuteurs pulaar. Bien que les élèves wolophones ayant bénéficié du traitement aient montré un effet négatif sur la compréhension orale, ils ont démontré des gains plus importants en matière de décodage (0,49 *écart-type* contre 0,33 *écart-type* pour les élèves pulaarophones) et ont obtenu un effet de traitement de 0,53 *écart-type* sur la fluidité de lecture à voix haute, alors que l'effet

de traitement pour les élèves pulaarophones en matière de fluidité de lecture à voix haute n'était pas statistiquement significatif.

3. Au niveau régional, les résultats les plus probants ont été obtenus à Kaolack. Dans cette région, trois des cinq sous-tests de lecture-écriture ont donné des résultats significativement positifs, contre un seul à Saint-Louis et aucun à Matam. Aucune des régions prises individuellement n'a obtenu d'effet positif et significatif dans la section consacrée aux mathématiques.

Nos conclusions s'inscrivent dans le contexte plus large des évaluations de l'éducation dans les PRFI. Une méta-analyse portant sur 32 pays (67 études) dans les PRFI a montré que les programmes pédagogiques structurés avaient un effet global de 0,30 *écart-type* sur les résultats en lecture-écriture fondamentale, avec des effets plus marqués sur les sous-compétences de reconnaissance des mots de niveau inférieur, et avec un impact variant en fonction du degré de soutien des enseignants et du fait que le programme était dispensé dans la langue maternelle (L1) ou dans une deuxième langue (L2) (Kim, Lee et Zuilkowski, 2020). Cette tendance à privilégier les sous-compétences de reconnaissance des mots bénéficiant des programmes d'apprentissage précoce de la lecture dans les PRFI est courante et se reflète également dans les conclusions de la présente étude. Dans un autre exemple, une évaluation d'impact portant sur une série d'interventions en Zambie a mis en évidence des effets positifs des effets positifs et significatifs compris entre 0,18 et 0,32 *écart-type* sur des sous-tests EGRA pour le décodage, la fluidité de lecture à voix haute et la compréhension en lecture, ainsi que des effets compris entre 0,13 et 0,14 *écart-type* sur les sous-tests EGMA sur le sens des nombres, l'addition et la soustraction (De Hoop et al., 2025). L'évaluation d'une intervention en lecture dans les premières années de scolarité au Libéria a rapporté des tailles d'effet allant de 0,52 à 1,23 *écart-type* pour la connaissance des sons des lettres, 0,55 pour la conscience phonémique, 0,78 pour le décodage de mots familiers, 1,23 *écart-type* pour le décodage des pseudo-mots, 0,80 *écart-type* pour la fluidité de lecture à voix haute et 0,82 *écart-type* pour la compréhension en lecture. La même étude a également révélé des tailles d'effet comprises entre 0,14 et 0,34 *écart-type* pour les sous-tests EGMA, bien que l'intervention ne comportait pas de composante mathématique (RTI, 2011). Enfin, une évaluation d'un programme d'enseignement de rattrapage en Colombie, qui utilisait EGRA pour mesurer les compétences en lecture-écriture des élèves, a révélé un impact de 0,36 *écart-type* sur la connaissance des sons des lettres et de 0,19 *écart-type* sur la fluidité de lecture à voix haute. Elle n'a révélé aucun impact significatif sur le décodage des pseudo-mots ou la compréhension en lecture (Álvarez Marinelli, 2024).

Enfin, nous estimons le coût total du *programme Ndaw Wune Année Scolaire* à **251 141 271 F CFA (433 002 USD)**, soit **83 714 F CFA (144 USD) par élève**. Nous constatons que la majeure partie du coût est consacrée aux activités d'enseignement direct (62 %), suivies par la gestion, la planification et l'évaluation (23 %). Environ la moitié des coûts du programme (51 %) sont consacrés aux frais de personnel directs, tandis que près d'un cinquième (20 %) des coûts sont utilisés pour financer les déplacements liés aux formations, aux journées pédagogiques et aux évaluations. **L'analyse coût-efficacité** révèle qu'un investissement de 100 dollars par élève produit une augmentation modérée de l'effet sur la connaissance des sons des lettres, le

décodage des mots et la fluidité de lecture à voix haute, ainsi qu’une légère augmentation de l’effet sur la compréhension en lecture et les résultats en mathématiques.³⁴

8. Recommandations pour accroître davantage l’impact

Enfin, nous présentons plusieurs recommandations à l’intention d’ARED, de la Fondation Gates et du MEN pour améliorer les programmes à l’avenir :

Pour le perfectionnement professionnel des enseignants et d’enseignement en classe :

1. **Élargir l’enseignement explicite de la phonétique aux classes ordinaires.** Le programme *Ndaw Wune* a démontré que le renforcement des compétences de mise en correspondance graphophonologique par un enseignement explicite de la phonétique, combiné à des opportunités structurées visant à améliorer l’automatisme (rapidité et précision dans l’établissement des correspondances graphème-phonème aux niveaux sous-lexical et lexical), s’avère particulièrement efficace pour améliorer les sous-compétences de décodage et la compréhension en lecture. Au regard des résultats nettement positifs observés pour ces sous-compétences dans le cadre de l’évaluation de *Ndaw Wune*, nous recommandons également que le **MOHEBS** s’appuie sur ces enseignements et envisage l’intégration systématique d’un enseignement explicite de la phonétique dans les langues sénégalaises dans les classes sur l’ensemble du territoire, pratique qui demeure absente dans la majorité des salles de classe (plus de 73 % selon Nakamura et al., 2024) au Sénégal.
2. **Veiller à l’expansion et à la mise à l’échelle des composantes de pédagogie structurée pour l’enseignement des compétences fondamentales dans les salles de classe à travers le Sénégal.** La combinaison d’une planification fondée sur une progression structurée des apprentissages (*scope and séquence*), de guides pédagogiques clairs et scénarisés (dans une certaine mesure), de supports d’enseignement et d’apprentissage, en particulier de cahiers d’exercices pour chaque enfant, a démontré, au regard des résultats de la présente évaluation, un impact significatif sur le développement des compétences ciblées et devrait être garantie dans les programmes de formation des enseignants et suivie par le biais d’un contrôle de la fidélité d’exécution dans les programmes d’apprentissage fondamental à travers le pays. Compte tenu du manque de cahiers d’exercices et de matériel d’enseignement et d’apprentissage dans les salles de classe, de l’exposition limitée à l’écrit en dehors du cadre scolaire, ainsi que de l’état souvent dégradé des tableaux noirs — difficiles à lire dans des salles de classe de grande taille — la mise à disposition d’un cahier individuel pour chaque élève constitue une composante essentielle des dispositifs de pédagogie structurée nécessaires à l’amélioration des apprentissages fondamentaux.
3. **Envisager un renforcement supplémentaire de l’appui pédagogique en compréhension orale dans les langues nationales.** Les résultats de l’évaluation n’ont mis en évidence aucune amélioration significative des compétences langagières

³⁴ Il existe peu d’analyses des coûts des programmes éducatifs en Afrique subsaharienne, et celles qui existent concernent généralement des programmes intégrés dans la structure éducative existante (c’est-à-dire que les coûts de personnel sont inexistantes ou très faibles). Il convient donc d’être prudent lorsque l’on compare le coût de la mise en œuvre du programme autonome *Ndaw Wune Année Scolaire* à celui d’autres programmes éducatifs.

orales, pourtant constitutives d'un des piliers essentiels du développement des compétences fondamentales en lecture-écriture. Par ailleurs, ce résultat ne saurait être attribuée à un effet plafond au niveau de la situation initiale. Par conséquent, *Ndaw Wune* et les autres programmes nationaux doivent renforcer la compréhension orale et les compétences lexicales orales, même s'ils enseignent déjà dans des langues que les enfants connaissent et comprennent. En d'autres termes, même lorsque les enfants reçoivent un enseignement dans leur propre langue, il est nécessaire de consacrer un temps d'enseignement considérable au renforcement du vocabulaire et des connaissances linguistiques dans l'ensemble du programme, et de ne pas se concentrer uniquement sur l'enseignement de la phonétique et du décodage.

4. **Investir dans l'amélioration des composantes du programme liées aux mathématiques.** L'évaluation d'impact a mis en évidence des difficultés dans l'enseignement des mathématiques. Nous proposons donc quelques approches pour améliorer l'enseignement des bases des mathématiques. Premièrement, l'enseignement devrait suivre une progression explicite au sein de chaque opération, en passant d'une perception des nombres à plusieurs chiffres comme une simple juxtaposition de chiffres isolés vers une compréhension structurée de la valeur positionnelle (par exemple, comprendre que 10 unités équivalent à une dizaine) (Rogers, 2023). Deuxièmement, les enseignants pourraient introduire explicitement des stratégies de calcul efficaces allant au-delà du simple comptage par dénombrement, telles que le passage par la dizaine, les sauts numériques, ou encore les méthodes d'arrondi et d'ajustement, dont l'efficacité pour améliorer la fluidité du calcul mental a été démontrée lorsqu'elles sont modélisées et pratiquées de manière systématique (Kumm & Graven, 2024). Troisièmement, l'intégration de discussions structurées autour des nombres permettrait d'encourager les élèves à expliciter et comparer différentes stratégies de résolution, favorisant ainsi le développement du sens du nombre, de la flexibilité cognitive et des compétences métacognitives grâce à un discours collaboratif. Enfin, le recours intensif à des supports pédagogiques de groupement, tel que des paquets de bâtonnets ou des blocs de base dix fabriqués localement, peut rendre tangibles des concepts abstraits liés à la valeur positionnelle (notamment l'idée que 10 unités constituent une dizaine) et faciliter le regroupement dans les opérations (Resnik, 1987).
5. **Nécessité de renforcer le transfert interlinguistique et les compétences linguistiques orales en français au regard des objectifs d'éducation bilingue du Sénégal.** L'objectif de *Ndaw Wune* est d'améliorer les compétences en lecture-écriture dans les langues nationales. L'objectif des apprentissages fondamentaux au Sénégal est l'acquisition, par les élèves, de compétences en lecture-écriture dans les langues nationales et le français (Nakamura et al., 2026). Les sciences de la lecture en contexte bilingue montrent clairement que l'apprentissage de la lecture dans une nouvelle langue est fortement tributaire de la capacité des élèves à lire dans leur langue maternelle et de leur niveau de maîtrise orale (ou signée) de la langue additionnelle à apprendre. Un programme de littératie fondamentale efficace au Sénégal devrait donc concevoir des curricula et des supports pédagogiques en tenant compte, dès le début du primaire, des compétences transférables et des fondements cognitifs de l'acquisition de la bilittératie pour faciliter la transition vers le primaire supérieur.

Pour des recherches plus approfondies :

- 1. Comprendre les interactions entre le développement du langage et le développement des compétences en mathématiques.** Les recherches futures ainsi que la conception des programmes au Sénégal devraient examiner de façon systématique la manière dont les compétences linguistiques et le choix de la langue d'enseignement influencent la progression des apprentissages en mathématiques. Les recherches devraient analyser dans quelle mesure l'enseignement bilingue ou multilingue, au sein duquel les élèves apprennent les mathématiques à la fois dans leur langue maternelle et en français, influe sur le raisonnement mathématique, le transfert de compétences et les résultats à long terme. Il conviendrait notamment d'évaluer si une maîtrise précoce des mathématiques en wolof ou en pulaar facilite ultérieurement l'apprentissage des mathématiques en français, ou si elle engendre des difficultés lors de la transition vers le vocabulaire et les concepts mathématiques. Les recherches futures devraient également étudier comment les inadéquations linguistiques affectent les résultats en mathématiques facilite ultérieurement l'apprentissage des mathématiques en français, ou si elle engendre, au contraire, des difficultés lors de la transition vers le vocabulaire et les concepts mathématiques.
- 2. Veiller à la collecte des données sur la correspondance entre la langue de l'enseignants.** Avec déploiement d'initiatives d'envergure en langues nationales, le MEN envisage de renforcer la collecte des informations sur la langue des enseignants au sein de MIRADOR (système d'information pour la gestion des enseignants), en appui à l'opérationnalisation de l'éducation bilingue. Nous recommandons de compléter ces données par des données linguistiques relatives aux élèves, de manière à mesurer l'ampleur des correspondances et des inadéquations entre les langues de l'enseignant et celles des élèves.
- 3. Réaliser une évaluation du processus pour approfondir les composantes du programme les plus susceptibles d'améliorer les acquis des élèves et d'expliquer les raisons des variations régionales.** Au moment où ARED cherche à affiner et à mettre à l'échelle le programme *Ndaw Wune*, une évaluation correspondante du processus serait extrêmement utile pour identifier les composantes clés du programme qui sont bénéfiques pour les élèves et les enseignants et pour préciser comment le programme peut reproduire les succès et résoudre les problèmes afin de l'optimiser en vue de son extension. Une telle évaluation pourrait analyser la fidélité de la mise en œuvre en comparant les activités prévues, telles que la prestation structurée des cours et l'utilisation des supports pédagogiques, avec ce qui a effectivement été réalisé dans les écoles. Elle pourrait également évaluer la qualité de la prestation au moyen d'observations en classe permettant de déterminer dans quelle mesure les enseignants appliquent les méthodes pédagogiques promues par le programme, la cohérence du coaching et la nature des adaptations opérées dans la pratique. Le fait de recueillir les points de vue des enseignants par le biais d'entretiens, d'enquêtes et de groupes de discussion pourrait aider à mettre en lumière la façon dont les enseignants perçoivent l'utilité de la formation et des supports pédagogiques, les défis auxquels ils sont confrontés et les conditions qui favorisent ou entravent le plus une

mise en œuvre efficace. L'évaluation pourrait également examiner de près l'engagement des élèves et la dynamique de classe pour identifier les pratiques pédagogiques qui conduisent le plus systématiquement à une participation active et à la compréhension. En outre, le suivi de la dose du programme (c'est-à-dire le temps que les élèves passent réellement à suivre les cours *Ndaw Wune*) pourrait contribuer à expliquer les variations d'effets observées selon les écoles et les régions.

4.

5. **Assurer un suivi systématique des données relatives à l'assiduité des élèves.** Étant donné que les données relatives à l'assiduité que nous avons reçues étaient incohérentes et présentaient plusieurs lacunes, les travaux futurs devraient envisager la mise en place d'un dispositif de suivi systématique et rentable permettant de suivre l'assiduité de chaque élève. Un tel dispositif permettrait d'identifier les causes des absences et de détecter l'absentéisme chronique. Il pourrait s'agir d'un simple outil basé sur des messages texte (SMS) qui alimenterait un dispositif de prise de décision fondé sur des données probantes et ne nécessiterait pas l'accès à Internet ou à un téléphone intelligent. Une telle approche a été testée avec succès dans plusieurs pays, notamment en Afrique de l'Ouest (Mugo & Piper, 2025). À titre d'exemple, un projet pilote au Kenya a montré que le coût de maintenance d'un système centralisé de suivi de l'assiduité par SMS pour 10 000 élèves au niveau central, complété par l'envoi hebdomadaire de SMS personnalisés aux enseignants concernant les élèves en difficulté, coûterait moins de 0,08 USD par enfant et par an (Mugo & Piper, 2025). L'amélioration des systèmes de données permettra également de suivre l'assiduité, l'attrition et les impacts hétérogènes, ce qui permettra une gestion plus réactive des programmes.

6. **Mobiliser l'intelligence artificielle (IA) et les technologies éducatives (le cas échéant) pour appuyer la production de supports d'enseignement et d'apprentissage (TLM) et les dispositifs de formation.** En matière de formation, de courtes vidéos générées par l'IA pourraient être utilisées pour des sessions fréquentes à faible intensité et pour un accompagnement de soutien. Elles peuvent également constituer un apport complémentaire sur des thématiques difficiles ou des séances pratiques liées aux mathématiques. Il peut également être utile d'utiliser l'IA pour produire des supports additionnels « de secours », conformes à tous les principes scientifiques de la lecture bilingue évoqués ci-dessus. À titre d'exemple, il peut s'agir de :

- a. veiller à chaque élève ait accès au manuel scolaire, que chaque enseignant dispose d'un guide et qu'ils soient alignés. Dans les rares cas où les élèves doivent partager des manuels, la formation des enseignants devrait intégrer des consignes pratiques sur l'animation d'activités de partage.
- b. recourir à l'IA pour créer des images utilisables comme cartes d'histoires, cartes de lettres et autres supports, de manière plus efficiente sur le plan des coûts, sans dépendre systématiquement des contraintes linguistiques actuelles de l'IA.
- c. veiller à ce que le contenu du manuel scolaire couvre toutes les sous-compétences en lecture-écriture et de langage selon une trajectoire

d'apprentissage appropriée—en particulier pour l'acquisition de la bilittératie et la transition ultérieure vers le français—afin de réduire la fragmentation des supports à imprimer et à harmoniser.

Pour les politiques :

- 7. Importance de commencer par enseigner aux enfants les langues nationales du Sénégal dans le cadre du MOHEBS.** Le programme *Ndaw Wune* montre clairement que lorsque l'enseignement est dispensé dans des langues que les enfants utilisent et comprennent, en mettant l'accent sur l'enseignement explicite des sons des lettres et des règles phonétiques, on constate des améliorations significatives dans les capacités de décodage et de compréhension en lecture, même chez les lecteurs les plus en difficulté. Il est important de noter que *Ndaw Wune* est un exemple exceptionnel qui montre que lorsque l'enseignement est dispensé dans des langues que les élèves comprennent, les ressources peuvent être concentrées sur la composante « reconnaissance des mots » de l'apprentissage fondamental de la lecture-écriture. Cette approche, par opposition à l'enseignement directement en français, rend l'acquisition de la lecture-écriture plus efficiente, un temps de classe précieux n'a pas à être mobilisé pour l'enseignement de compétences orales de base et peut être consacré prioritairement aux sons des lettres, à l'assemblage, à la fluidité et au développement de compétences orales de niveau supérieur. En plus du fait que les élèves comprennent la langue d'enseignement, il est également indispensable que la langue de l'enseignant corresponde à celle des élèves. Les données de la présente étude pas mis en évidence de différences significatives entre la langue des élèves et la langue d'enseignement. Toutefois, nous n'avons pas collecté de données sur les compétences linguistiques des enseignants. Des inadéquations importantes persistent encore à l'échelle nationale entre la langue des élèves, celle des enseignants et la langue d'enseignement (Nakamura et al., 2024 ; Diallo & Mbaye, 2025) et doivent être corrigées afin d'améliorer les apprentissages fondamentaux, tant en langues nationales qu'en français.

Références

- Ahouakan Ehouman, W. V., & Diene, M. (2017). Does school quality matter? Primary schools' characteristics and child labor intensity in Senegal. *Journal of African Development*, 19(1), 113–131. <https://doi.org/10.5325/jafrideve.19.1.0113>
- All children in Senegal are born to learn.* (n.d.). https://www.unesco.org/gem-report/sites/default/files/medias/fichiers/2023/02/Spotlight%20Advocacy_Senegal_EN_v1.pdf
- Álvarez Marinelli, H. A., Berlinski, S., & Busso, M. (2024). Remedial education: Evidence from a sequence of experiments in Colombia. *Journal of Human Resources*, 59(1), 141–174. <https://doi.org/10.3368/jhr.0320-10801R2>
- Angrist, N., Meager, R., Angrist, & London School of Economics. (2023). Implementation matters: Generalizing treatment effects in education. In *EdWorkingPaper* (No. 23–802). <https://edworkingpapers.com/sites/default/files/ai23-802.pdf>
- Bagby, E., Swift-Morgan, J., Niang, A., Upadhyay, A., & Chemonics International Inc. (2022). Achieving Cost-Effective Instructional Coaching at Scale: Evidence from Senegal. In *Learning From All Children Reading*. https://chemonics.com/wp-content/uploads/2022/04/ACR_Tech_Brief_LPT_Coaching_CIES.pdf
- Banerjee, A. V., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., & Walton, M. (2016, November 20). *Mainstreaming an Effective intervention: Evidence from Randomized randomized evaluations of “Teaching at the Right Level” in India.* (MIT Department of Economics Working Paper No. 16-08). SSRN. <https://doi.org/10.2139/ssrn.2846971>
- Brander, R. L., Puett, C., Becquey, E., Leroy, J. L., Ruel, M. T., Sessou, F. E., & Huybregts, L. (2024). The cost and cost-effectiveness of an integrated wasting prevention and screening intervention package in Burkina Faso and Mali. *The Journal of Nutrition*, 154(8), 2551–2565. <https://doi.org/10.1016/j.tjnut.2024.04.010>
- Bronfenbrenner, U. (2000). Ecological systems theory. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology* (Vol. 3, pp. 129–133). Oxford University Press. <https://doi.org/10.1037/10518-046>

Brown, L. E., Kim, H. Y., Dolan, C. T., Brown, A., Sklar, J., & Aber, J. L. (2022). Remedial Programming and Skill-Targeted SEL in Low-Income and Crisis-Affected contexts: Experimental evidence from Niger. *Journal of Research on Educational Effectiveness*, 16(4), 583–614. <https://doi.org/10.1080/19345747.2022.2139785>

Cost-effective approaches cost-effective approaches to improve global to improve global learning. (2023). *What does recent evidence tell us are “Smart Buys” for improving learning in low-and middle-income countries? Recommendations of the Global Education Evidence Advisory Panel (GEEAP)*. <https://thedocs.worldbank.org/en/doc/231d98251cf326922518be0cbe306fdc-0200022023/related/GEEAP-Report-Smart-Buys-2023-final.pdf>

Danks, A., & Kendziora, K. (2021, April). *Standards for the economic evaluation of educational and social programs*. American Institutes for Research. <https://www.air.org/sites/default/files/Standards-for-the-Economic-Evaluation-of-Educational-and-Social-Programs-CASP-May-2021.pdf>

De Hoop, T., Ring, H., Siwach, G., Dias, P., Tembo, G., Taube, A., & Behmer, T. (2025, January). *Mixed-methods cluster-randomized controlled trial of Impact Network’s eSchool 360 model in rural Zambia: Endline report*. American Institutes for Research. <https://www.air.org/sites/default/files/2025-01/Endline-Report-Impact-Network-eSchool-RCT-Trial-360-Zambia-Jan-2025.pdf>

Duflo, A., Kiessel, J., & Lucas, A. (2020). *Experimental evidence on alternative policies to increase learning at scale*. <https://doi.org/10.3386/w27298>

Duflo, Annie, et al. “Experimental Evidence on Four Policies to Increase Learning at Scale.” *the Economic Journal/Economic Journal*, 2 Feb. 2024, <https://doi.org/10.1093/ej/ueae003>.

Encinar, Pierre-Emmanuel. (2022). PASEC 2019 study report in English

Evans, D., & Mendez Acosta, A. (2020). *Education in Africa: What Are We Learning?* <https://www.cgdev.org/sites/default/files/education-africa-what-are-we-learning.pdf>

Evans, D. K., & Mendez Acosta, A. (2021). Education in Africa: What are we learning? *Journal of African Economies*, 30(1), 13–54. <https://doi.org/10.1093/jae/ejaa009>

Fleisch, B., Taylor, S., Schöer, V., & Mabogoane, T. (2017). Failing to catch up in reading in the middle years: The findings of the impact evaluation of the Reading Catch-Up Programme in South Africa. *International Journal of Educational Development*, 53, 36–47. <https://doi.org/10.1016/j.ijedudev.2016.11.008>

- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and psychological measurement*, 33(3), 613–619.
- Fuson, K. C., & Briars, D. J. (1990). Using a Base-Ten Blocks Learning/Teaching Approach for First- and Second-Grade Place-Value and Multidigit Addition and Subtraction. *Journal for Research in Mathematics Education*, 21(3), 180. <https://doi.org/10.2307/749373>
- Greenwood, C. R., Carta, J. J., Walker, D., Watson-Thompson, J., Gilkerson, J., Larson, A. L., & Schnitz, A. (2017). Conceptualizing a public health prevention intervention for bridging the 30 million word gap. *Clinical child and family psychology review*, 20(1), 3-24.
- Harden, K., Punjabi, M., & Fernandez, M. (2022). Influences on Teachers' Use of the Prescribed Language of Instruction: Evidence from Four Language Groups in the Philippines. *Education Quarterly Reviews*, 5(1). <https://doi.org/10.31014/aior.1993.05.01.460>
- Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American educator*, 27(1), 4-9.
- Hoover, W. A., & Tunmer, W. E. (2020). *The cognitive foundations of reading and its acquisition: A framework with applications connecting teaching and learning*. Springer International Publishing. <https://link.springer.com/book/10.1007/978-3-030-44195-1>
- Jacob, B. A., & Lefgren, L. (2002). *The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago* (NBER Working Paper No. 8916). National Bureau of Economic Research. <https://www.nber.org/papers/w8916>
- Kébé, A. (n.d.). *Revue ASTR-GHANA Language and Multilingualism in Senegal: From Colonial Legacies to Contemporary Educational Issues*. Retrieved October 1, 2025, from <https://edition-efua.acaref.net/wp-content/uploads/sites/6/2025/08/1-Abou-Bakry-KEBE.pdf>
- Koda, K. (2008). Impacts of prior literacy experience on learning to read in a second language. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 68–96). Routledge.

- Kumm, M., & Graven, M. H. (2024). Exploring pre-service teachers' knowledge of efficient calculation strategies. *South African Journal of Childhood Education*, 14(1).
<https://doi.org/10.4102/sajce.v14i1.1494>
- LARTES-IFAN. (2017). Jàngandoo, Baromètre de la qualité des apprentissages au Sénégal - Principaux résultats.
- Levin, H. M., McEwan, P. J., Belfield, C. R., Bowden, A. B., & Shand, R. D. (2018). *Economic evaluation in education: Cost-effectiveness and benefit-cost analysis* (3rd ed.). SAGE Publications. <https://searchworks-lb.stanford.edu/view/13785876>
- Maruyama, T., & Kurosaki, T. (2021). Do remedial activities using math workbooks improve student learning? Empirical evidence from scaled-up interventions in Niger. *World Development*, 148, Article 105659. <https://doi.org/10.1016/j.worlddev.2021.105659>
- Maruyama, Takao and Igei, Kengo. (2022). Community-Wide Support for Primary Students to Improve Basic Reading and Math Learning: Empirical Evidence from Madagascar
<https://ssrn.com/abstract=4076787> or <http://dx.doi.org/10.2139/ssrn.4076787>
- Mohohlwane, N., Taylor, S., Cilliers, J., & Fleisch, B. (2024). Reading skills transfer best from home language to a second language: Policy lessons from two field experiments in South Africa. *Journal of Research on Educational Effectiveness*, 17(4), 687–710.
<https://doi-org.air.idm.oclc.org/10.1080/19345747.2023.2279123>
- Monitoring Impacts on Learning Outcomes MAIN REPORT*. (n.d.).
<https://milo.uis.unesco.org/wp-content/uploads/sites/17/2022/01/MILO-Summary-Full-Report.pdf>
- Moussa, S., Cissé, R., and Lô, C. (2021). How remediation helps to improve children's educational learnings in primary education in Senegal. In I. I. Munene (Ed.), *Ensuring all children learn: Lessons from the South on what works in equity and inclusion*. Lexington Books.
- Mugo, P., & Piper, B. (2025, January 14). *Using SMS as a low-cost means to systematically track chronic absenteeism in schools: Reflections from South Sudan and Sierra Leone*. UKFIET – The Education and Development Forum.
<https://www.ukfiet.org/2025/using-sms-as-a-low-cost-means-to-systematically-track-chronic-absenteeism-in-schools-reflections-from-south-sudan-and-sierra-leone/>
- Muyabi, P. (2017). *Parental involvement in reducing learner absenteeism in selected primary schools of Kazungula district in the Southern Province of Zambia*.
<https://dspace.unza.zm/items/0eb43fa2-469b-4b03-beee-60fb3078e940>

- Nag, S., Vagh, S. B., Dulay, K. M., & Snowling, M. J. (2019). Home language, school language and children's literacy attainments: A systematic review of evidence from low- and middle-income countries. *Review of Education*, 7(1), 91–150.
<https://doi.org/10.1002/rev3.3130>
- Nakamura, P., Molotsky, A., Zarzur, R. C., Ranjit, V., Haddad, Y., & De Hoop, T. (2023). Language of instruction in schools in low-and middle-income countries: A systematic review. *Campbell Systematic Reviews*, 19(4), Article e1351.
<https://doi.org/10.1002/cl2.1351>
- Nakamura, P. R. & Holla, C. U. (2024). Semantic fluency: Measuring oral language proficiency to enable informed decision making on Language of Instruction in six countries. Comparative and International Education Society (CIES). Miami, Florida.
- Nakamura, P., McBride, A., Holla, C., Yessifou, N., & Woodrick, K. (2026). *Blueprint for the use of evidence in language of instruction (LoI) decisions (BUILD): Senegal report* (Unpublished report). Bill & Melinda Gates Foundation.
- Nakamura, P., Woodrick, K., McBride, A., Holla, C., & Zaas, D. (n.d.). *Blueprint for the Use of Evidence in Language of Instruction (LoI) Decisions (BUILD) Framework Introduction*. Retrieved January 27, 2026, from <https://www.air.org/sites/default/files/2025-08/BUILD-Introduction-to-the-Framework-August-2025.pdf>
- Ndaw Wune: An Accelerated Learning Program to Support Education in Senegal - ARED. (2024, November 21). Ared-Edu.org. <https://ared-edu.org/en/ndaw-wune-an-accelerated-learning-program-to-support-senegalese-entrepreneurship/>
- Ndour, C. T. (2024). *Understanding school dropout in Senegal: Exploring the impact of governance factors* (SSRN Scholarly Paper No. 4867074). Social Science Research Network. <https://doi.org/10.2139/ssrn.4867074>
- Nishanthi, R. (2020). Understanding of the importance of mother tongue learning. *International Journal of Trend in Scientific Research and Development*, 5(1), 77–80.
<https://www.ijtsrd.com/papers/ijtsrd35846.pdf>
- Ouellette, G., Beers, A. A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Read Writ* 23, 189–208 (2010).
<https://doi.org/10.1007/s11145-008-9159-1>
- Özek, U. (2021). *The effects of middle school remediation on postsecondary success: Regression discontinuity evidence from Florida* (Working Paper No. 258-0921). National Center for Analysis of Longitudinal Data in Education Research (CALDER).
<https://eric.ed.gov/?id=ED616057>

- Pillai, S. (2025, June 23). *Using SMS as a low cost means to systematically track chronic absenteeism in schools: Reflections from South Sudan and Sierra Leone*. The Education and Development Forum (UKFIET). <https://www.ukfiet.org/2025/using-sms-as-a-low-cost-means-to-systematically-track-chronic-absenteeism-in-schools-reflections-from-south-sudan-and-sierra-leone/>
- Piper, B., Ralaingita W., Mejia J., Dubeck, M. M., DeStefano J., Stern J., Jordan R., & Sitabkhan Y. (2020). Structured Pedagogy How-to Guides and Literature Review https://scienceofteaching.site/wp-content/uploads/2022/10/SP_Introduction.pdf
- Piper, B. & Dubeck, M. M. (2024). Responding to the learning crisis: Structured pedagogy in sub-Saharan Africa. *International Journal of Educational Development*, 109, Article 103095. <https://doi.org/10.1016/j.ijedudev.2024.103095>
- Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, 16(9), 13–20. <https://www.jstor.org/stable/749373>
- Rogers, A. (2023). *Rethinking the number magnitude-based progression: An analysis of place value development in Years 3–6* (MERGA45 Proceedings). Mathematics Education Research Group of Australasia. <https://files.eric.ed.gov/fulltext/ED631611.pdf>
- RTI International. (2011). *Information for accountability: Impact evaluation of EGRA and teacher training (English)* (Working Paper No. 72457). World Bank. <http://documents.worldbank.org/curated/en/772611468057237703>
- Singh, A., Mauricio Romero, Muralidharan, K., Alison Andrew, Noam Angrist, Jishnu Das, Jean Dreze, Peter Hull, Ajinkya Kesar, Ben Piper, Justin Sandefur, Ramamurthy Sripada, Dwithiya Raghavan, Akanksha Soni, & Indira Tirumala. (2024). COVID-19 Learning loss and recovery: Panel data evidence from India. In *INST TECH AUTONOMO DE MEXICO* [Journal-article]. <https://doi.org/10.3368/jhr.0723-13025R2>
- Snilstveit, B., Stevenson, J., Phillips, D., Vojtkova, M., Gallagher, E., Schmidt, T., Jobse, H., Geelen, M., Pastorello, M., and Eyers, J. (2015, December). *Interventions for improving learning outcomes and access to education in low- and middle- income countries: A systematic review* (3ie Systematic Review 24). International Initiative for Impact Evaluation. <https://doi.org/10.23846/SRS007>
- Stern, J., Jukes, M., Piper, B., DeStefano, J., Mejia, J., Dubeck, P., Carrol, B., Jordan, R., Gatuyu, C., Nduku, T., Punjabi, M., Harris Van Keuren, C., & Tufail, F. (2021, October). *Learning at scale: Interim report*. RTI International. <https://ierc-publicfiles.s3.amazonaws.com/public/resources/Learning%20at%20Scale%20Interim%20Report%20-%20Final%20Draft.pdf>

- Stern, J., Jukes, M. C. H., Cilliers, J., Fleisch, B., Taylor, S., & Mohohlwane, N. (2024). Persistence and emergence of literacy skills: Long-term impacts of an effective early grade reading intervention in South Africa. *Journal of Research on Educational Effectiveness*. <https://doi.org/10.1080/19345747.2024.2417288>
- Nakamura, P., Molotsky, A., Holla, C., McBride, A., & Gosselt, J. (2024). *Strengthening Bilingual and Multilingual Learning Systems in Francophone Africa Evidence from Senegal*. (n.d.). <https://www.air.org/sites/default/files/2024-05/KIX-Strengthening-Bilingual-Education-Senegal-Brief-May-2024.pdf>
- Schwartz, A. (2012). *Remedial Education Programs to Accelerate Learning for All*. <https://documents1.worldbank.org/curated/en/564671468151507990/pdf/797830WP0Anali0Box0379789B00PUBLIC0.pdf>
- Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., Jobse, H., Schmidt, T., & Jimenez, E. (2016). *Systematic Review Summary 7 The impact of education programmes on learning and school participation in low-and middle-income countries*. <https://www.3ieimpact.org/sites/default/files/2019-05/srs7-education-report.pdf>
- Snilstveit, B., Gallagher, E., Phillips, D., Vojtkova, M., Eyers, J., Skaldiou, D., Stevenson, J., Bhavsar, A., & Davies, P. (2017). PROTOCOL: Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. *Campbell Systematic Reviews*, 13(1), 1–82. <https://doi.org/10.1002/cl2.176>
- The Abdul Latif Jameel Poverty Action Lab (J-PAL)*. (n.d.). The Abdul Latif Jameel Poverty Action Lab (J-PAL). <https://www.povertyactionlab.org/evaluation/impact-government-run-remedial-education-programs-student-learning-ghana>
- UNESCO Institute for Statistics. (2024). *Bulk Data Download Service*. <https://uis.unesco.org/bdds>
- UNICEF. (2022). Teacher attendance and time on task in West and Central Africa. <https://www.unicef.org/innocenti/media/4296/file/UNICEF-Time-to-Teach-Teacher-Attendance-WCA-2022.pdf>
- United States Agency for International Development (USAID). (2022). *Renforcement de la Lecture Initiale pour Tous (RELIT)*. https://www.rti.org/sites/default/files/relit_dec22_final_eng_approved.pdf

What Works Clearinghouse™ Standards Handbook Version 4.0. (n.d.).

https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

World Bank. (2024). *Senegal learning poverty brief - 2024 (English)*. World Bank Group.

<http://documents.worldbank.org/curated/en/099062624135013503>

Annexe A. Test de sélection d'échantillons

Test Elève en Pulaar

Lecture

a	i	n	t	l	r	m	w	d	k
Demander à l'élève de lire les sons (pas les noms) des lettres. Merci de reporter la note de l'élève sur 10 à la feuille de notation									
na	yo	fof	so	mo	to	wi	ko		
tan	ne								
Demander à l'élève de lire les syllabes. Merci de reporter la note de l'élève sur 10 à la feuille de notation									
taro	deftere			wiyi		nana			
wuro									
ladde	gooto			galle					
kala	woni								
Demander à l'élève de lire les mots. Merci de reporter la note de l'élève sur 10 à la feuille de notation									

Mathématiques

11 10

19 32

7 5

16 13

18 29

Demander à l'élève de vous dire le chiffre le plus grand. Merci de reporter la note de l'élève sur 5

Annexe B. Lieux de mise en œuvre du programme

Année scolaire (2024–2025)		Vacances (2024)	
IA	IEF	IA	IEF ³⁵
Diourbel	Bambey	Diourbel	- Bambey - Diourbel - Mbacké
Kaolack	Kaolack	Kaolack	Nioro du Rip
Saint-Louis	- Saint-Louis - Podor	Matam	Matam
Matam	- Matam - Kanel		

Annexe C. Méthodologie détaillée de calcul des coûts

Dans cette annexe, nous fournissons des renseignements supplémentaires concernant notre méthodologie de calcul des coûts et mettons en évidence les hypothèses que nous avons formulées concernant les analyses de coûts. Les pages suivantes contiennent des tableaux répertoriant les coûts inclus de manière plus détaillée.

Coûts de personnel d’ARED. ARED a fourni les données salariales de 16 employés, ainsi que le pourcentage de temps consacré à *Ndaw Wune Année Scolaire* pendant les neuf mois de l’année scolaire. Nous avons réparti les coûts de personnel d’ARED entre les différentes composantes du programme au prorata de la durée de chaque composante et de la participation des agents aux activités correspondantes. Par exemple, la mise en œuvre directe de *Ndaw Wune Année Scolaire* s’étend sur six des neuf mois, tandis que la composante gestion, planification et évaluation couvre les trois mois restants ; par conséquent, environ **67 %** des coûts de personnel d’ARED sont affectés à la composante mise en œuvre directe et 33 % à la composante gestion, planification et évaluation. Pour le personnel participant à la formation initiale des enseignants et/ou aux journées pédagogiques, nous allouons 2,56 % du salaire hebdomadaire (en supposant un salaire de neuf mois couvrant 39 semaines). Par exemple, un membre du personnel d’ARED qui a participé à la formation initiale des

³⁵ Il s’agit de six communes : Thilogne et Ourossogui à Matam, Ngogom à Bambey, Diourbel, Kael à Mbacké et Paoskoto à Nioro.

enseignants et à toutes les journées pédagogiques verrait son salaire annuel *Ndaw Wune Année Scolaire* réparti entre chaque composante de la manière suivante :

- 2,6 % pour la formation initiale des enseignants
- 59 % pour l'enseignement direct
- 5,1 % pour les journées pédagogiques
- 33,3 % pour la gestion, la planification et l'évaluation

Il convient de noter que les allocations totalisent 100 % pour l'ensemble du personnel d'ARED (les coûts totaux de personnel sont donc inclus dans les estimations globales des coûts du programme) et que la répartition du temps de travail du personnel par activité fournit une estimation approximative permettant de démontrer la répartition des coûts de personnel pour chaque activité.

Coûts de personnel non liés à ARED. ARED a engagé deux consultants pour la formation initiale des enseignants et un consultant pour les journées pédagogiques ; les consultants ont été rémunérés 60 000 F CFA par jour. Les enseignants, les superviseurs et les inspecteurs ont reçu une allocation pour les six mois de mise en œuvre du programme : les enseignants ont reçu 50 000 F CFA par mois (pour un total de 300 000 F CFA), les superviseurs ont reçu 100 000 F CFA par mois (pour un total de 600 000 F CFA) et les inspecteurs ont reçu 25 000 F CFA par trimestre (pour un total de 50 000 F CFA). Les inspecteurs ont reçu une allocation supplémentaire de 10 000 par jour pendant la formation des enseignants et les journées pédagogiques. Les évaluateurs ont reçu une allocation de 30 000 par personne et par jour de collecte de données (pour deux cycles de deux jours au total).

Matériel et équipement. Pour les coûts des supports pédagogiques, nous avons retenu les dépenses d'impression et de livraison engagées par ARED pour *Ndaw Wune Année Scolaire*. S'agissant des ordinateurs portables et des vidéoprojecteurs, un exemplaire de chaque équipement a été mobilisé dans chaque IEF pour les besoins des formations. Nous avons recueilli des prix moyens sur le marché local et estimé un coût journalier sur la base d'une durée de vie utile supposée de six ans pour les vidéoprojecteurs et de quatre ans pour les ordinateurs portables. Le coût journalier ainsi calculé a ensuite été ventilé entre la formation des enseignants et les journées pédagogiques, au prorata du nombre de jours d'utilisation.

Frais de déplacement. Nous nous sommes appuyés sur des échanges avec ARED, des documents budgétaires, les termes de référence (TDR) et le protocole G50 pour imputer les coûts de transport, d'hébergement et de restauration.

S'agissant des coûts de transport du personnel de ARED pendant la formation initiale, les journées pédagogiques et les tournées d'évaluation, nous avons suivi le calendrier de déplacement figurant dans les TDR des journées pédagogiques. Les TDR répartissent les IEF en trois axes, chaque axe comprenant deux IEF. Pour chaque axe, nous avons retenu l'hypothèse d'un chauffeur et d'un véhicule, conformément à l'itinéraire décrit dans les TDR (trajet vers la première IEF pour y déposer le personnel, puis vers la deuxième, retour vers la première pour récupérer le personnel, avant retour à Dakar). Nous avons supposé un forfait journalier total de 95 000 FCFA par binôme chauffeur/véhicule, incluant la rémunération du chauffeur, les repas et l'hébergement. Nous avons également supposé une consommation de 8,5 km par litre et un coût du carburant de 755 FCFA par litre.

Pour le transport des personnes ne relevant pas de ARED entre régions, nous avons calculé les remboursements sur la base de 75 FCFA par kilomètre, en fonction des distances entre villes, conformément au protocole G50 et aux informations budgétaires de ARED. Pour les déplacements au sein des régions, nous avons estimé des indemnités journalières distinctes selon que le participant réside dans le village où se déroule l'événement ou qu'il se déplace depuis un autre lieu, en posant l'hypothèse que, pour les événements organisés dans les écoles, un enseignant et un directeur seraient résidents du village et que l'ensemble des autres participants viendraient d'autres localités de la région. Les taux locaux retenus étaient de 5 000 FCFA pour les résidents et de 10 000 FCFA pour les non-résidents. Pour les journées pédagogiques, nous avons retenu une indemnité de **10 000 FCFA** pour l'ensemble des participants, conformément aux échanges avec ARED.

S'agissant de l'hébergement, nous avons appliqué le montant du protocole G50, soit 30 000 FCFA par nuit, et n'avons comptabilisé ces coûts que pour les participants venant d'une autre région que celle où se tenait l'événement. Nous avons par ailleurs supposé que le personnel de ARED se déplaçait la veille et rentrait le lendemain lorsque l'événement se déroulait dans une autre région.

Pour la restauration, nous avons calculé des taux de remboursement pour l'ensemble des personnes en déplacement, selon qu'elles étaient résidentes de la région (auquel cas elles percevaient 6 000 FCFA par jour pour deux repas, conformément au protocole G50) ou non (auquel cas elles percevaient 10 000 FCFA par jour pour trois repas).

Frais généraux. Après consultation avec ARED au sujet des frais généraux, nous avons appliqué un taux forfaitaire de **10 %** à chaque activité.

Formation initiale des enseignants. Pour la formation initiale des enseignants, nous nous sommes appuyés sur les TDR de *Ndaw Wune Année Scolaire*, complétés par des informations issues d'entretiens avec le personnel de l'ARED. Les TDR et le document budgétaire fourni prévoyaient la participation à la journée d'harmonisation précédant la formation de cinq agents de l'ARED, deux consultants rémunérés, quatre facilitateurs additionnels, six enseignants modèles, dix superviseurs et dix points focaux/inspecteurs. Pour les quatre jours de formation, nous avons pris en compte les personnes susmentionnées, auxquelles s'ajoutaient 100 enseignants et 66 directeurs (ces derniers ne participant qu'au dernier jour).

Journées pédagogiques. Pour l'estimation des coûts des journées pédagogiques, nous nous sommes appuyés sur les termes de référence de *Ndaw Wune Année Scolaire*. Pour les indemnités journalières locales, nous avons supposé qu'un enseignant, un directeur et un superviseur résidaient dans le village où se déroulait l'événement pour chaque IEF, les autres résidant hors du village. Les TDR de *Ndaw Wune Année Scolaire* prévoyaient la participation de trois agents de l'ARED, un consultant rémunéré, trois facilitateurs additionnels, 100 enseignants, 10 superviseurs, 12 points focaux IEF, quatre points focaux IA et 56 directeurs.

Gestion, planification, suivi et évaluation. Pour le programme Année Scolaire, nous avons retenu l'hypothèse de deux tournées d'évaluation par an. Les coûts annuels incluent une journée de formation des évaluateurs (à laquelle participent quatre évaluateurs par IEF) et deux jours de collecte de données par tournée (réalisée par deux évaluateurs par IEF). Nous avons intégré les coûts annualisés des tablettes nécessaires à la collecte de données, imputés au prorata des jours d'utilisation, sur la base d'une durée de vie utile supposée de quatre ans

par tablette. Nous avons également intégré les coûts annualisés du logiciel SurveyCTO, que le responsable du suivi-évaluation de l'ARED a indiqué que l'organisation envisageait d'utiliser à l'avenir..

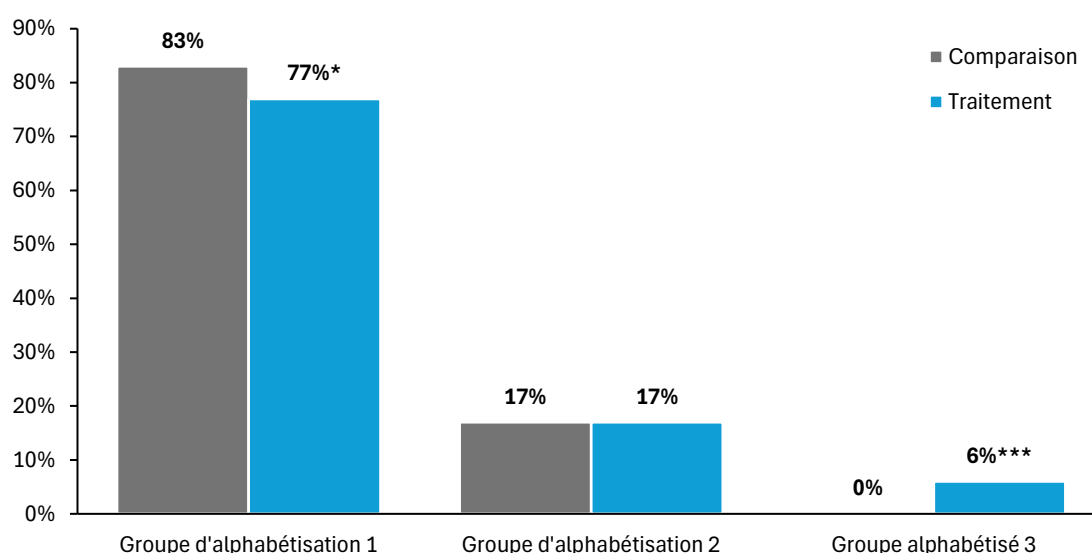
Annexe C1. Résumé des coûts annuels pour *Ndaw Wune Année Scolaire*

Catégorie de ressources	Coût (F CFA)	Coût (USD)
Formation initiale des enseignants		
Personnel	2 478 417	4 273
Personnel ARED	1 238 417	2 135
Personnel non ARED	1 240 000	2 138
Dépenses hors personnel	15 402 471	26 556
Déplacements	15 381 557	26 520
Équipement/Matériel	20 913	36
Installations	0	0
Frais généraux	1 986 765	3 425
Total des activités	19 867 653	34 255
Enseignement direct		
Personnel	92 867 258	160 116
Personnel ARED	56 567 258	97 530
Personnel non ARED	36 300 000	62 586
Dépenses hors personnel	47 037 704	81 099
Déplacements	0	0
Équipement/Matériel	47 037 704	81 099
Installations	0	0
Frais généraux	15 544 996	26 802
Total des activités	155 449 958	268 017
Journées pédagogiques		
Personnel	2 061 988	3 555
Personnel ARED	981 988	1 693
Personnel non ARED	1 080 000	1 862
Hors personnel	14 780 177	25 483
Déplacements	14 759 264	25 447
Équipement/Matériel	20 913	36
Installations	0	0
Frais généraux	1 871 352	3 226

Catégorie de ressources	Coût (F CFA)	Coût (USD)
Total des activités	18 713 517	32 265
Gestion, planification et évaluation		
Personnel	30 657 927	52 858
Personnel ARED	29 217 927	50 376
Personnel non ARED	1 440 000	2 483
Dépenses hors personnel	20 741 202	35 761
Déplacements	20 564 279	35 456
Équipement/matériel	151 923	262
Installations	25 000	43
Frais généraux	5 711 014	9 847
Total des activités	57 110 144	98 466
TOTAL	251 141 271	433 2

Annexe D. Tests de robustesse

Annexe D133 . Sélection dans les groupes de lecture-écriture *Ndaw Wune* au début de l'étude



Remarque : Dalberg a utilisé le même test de sélection que celui utilisé par ARED dans le programme *Ndaw Wune* pour déterminer les élèves éligibles. Les élèves du groupe de lecture-écriture 1 lisent moins de 10 lettres. Les élèves du groupe de lecture-écriture 2 lisent 10 lettres mais moins de 10 syllabes. Les élèves du groupe de

lecture-écriture 3 lisent 10 lettres et 10 syllabes. $N = 561$ élèves du groupe de comparaison et 405 élèves du groupe de traitement.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D234 . Équivalence de base entre les élèves perdus de vue et les élèves appariés

Résultat	Comparaison perdus de vue	Équivalence de base des élèves appariés	Différence	Traitement perdus de vue	Traitement appariés	Différence
Fille	65	47	-19 %***	51	50	-2
Âge	9,55 ans	9,49 ans	-0,06 an	9,52 ans	9,69 ans	0,17 an
Classe de CE1	97	95	-2	93	97	4
A fréquenté une école maternelle	64	56	-7	62	59	-3
LE : wolof	70	65	-5	0,43	0,3	-0,13*
LE : sérère	2	3	1	0,35	0,35	0,01
LE : pulaar	28	32	4	0,23	0,35	0,12**
Nourriture consommée aujourd'hui	81	83	3	81	82	1
A reçu des cours particuliers en dehors de l'école	70	78	7	88	88	0
Au cours de la dernière semaine complète d'école, combien de jours avez-vous manqué ?	1,56 jour	1,56 jour	0 jour	1,57 jour	1,31 jour	-0,26 jour**
Compréhension orale (sur 6)	3,25	3,11	-0,14	3,36	3,42	0,06
Conscience phonologique (sur 10)	7,83	7,53	-0,29	8,68	8,39	-0,28
Connaissance des sons des lettres (sur 100)	72,04	76,15	4,1	69,25	71,25	2
Décodage de pseudo- mots (sur 50)	6,13	6,78	0,65	5,58	6,12	0,54
Décodage de mots familiers (sur 50)	5,39	6,29	0,9	5,99	6,99	1
Fluidité de lecture à voix haute (sur 50)	4,09	5,53	1,45	5,35	5,64	0,29
Compréhension écrite (sur 6)	0,48	0,56	0,08	0,56	0,72	0,16

Résultat	Comparaison perdus de vue	Équivalence de base des élèves appariés	Différence	Traitement perdus de vue	Traitement appariés	Différence
Numéro d'identification (sur 24)	17,64	18,18	0,54	17,82	18,36	0,54
Comptage (sur 6)	5,18	5,09	-0,09	5,37	5,32	-0,05
Discrimination des chiffres (sur 12)	9,83	9,71	-0,11	10,37	10,4	0,04
Sens des nombres (sur 12)	4,43	4,43	-0,01	4,69	4,41	-0,28
Addition (sur 12)	6,03	6,25	0,22	6,71	7,11	0,4
Soustraction (sur 12)	4,4	4,86	0,46	4,78	5,23	0,45

Remarque : LE = langue d'enseignement. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D335 . Équivalence de base pour un échantillon apparié

Résultat	Groupe de comparaison	Groupe de traitement	Différence
Filles	47	50	3 %
Âge	9,49 ans	9,69 ans	0,19 an
Classe CE1	95	97	2
A fréquenté une école maternelle	56	59	3
LE : wolof	65	30	-35 %***
LE : sérère	3	35	32 %***
LE : pulaar	32	35	3
Nourriture consommée aujourd'hui	83	82	-1
A reçu des cours particuliers en dehors de l'école	78	88	11
Au cours de la dernière semaine complète d'école, combien de jours avez-vous manqué ?	1,56 jour	1,31 jour	-0,24 jour***
Compréhension orale (sur 6)	3,11	3,42	0,31
Conscience phonologique (sur 10)	7,53	8,39	0,86
Connaissance des sons des lettres (sur 100)	76,15	71,25	-4,9*
Décodage de pseudo-mots (sur 50)	6,78	6,12	-0,65
Décodage de mots familiers (sur 50)	6,29	6,99	0,70
Fluidité de lecture à voix haute (sur 50)	5,53	5,64	0,11

Résultat	Groupe de comparaison	Groupe de traitement	Différence
Compréhension écrite (sur 6)	0,56	0,72	0,16
Identification des nombres (sur 24)	18,18	18,36	0,18
Comptage (sur 6)	5,09	5,32	0,23
Discrimination des chiffres (sur 12)	9,71	10,4	0,69
Sens des nombres (sur 12)	4,43	4,41	-0,02
Addition (sur 12)	6,25	7,11	0,86
Soustraction (sur 12)	4,86	5,23	0,37

Remarque : LE = langue d'enseignement. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D436 . Équivalence de base sans les élèves sérères

Résultat	Groupe de comparaison	Groupe de traitement	Différence
Filles	50	51	0
Âge	9,51 ans	9,59 ans	0,07 an
Niveau CE1	95	94	-1
A fréquenté une école maternelle	59	58	-1
LE : wolof	68	51	-17
LE : sérère	0	0	0
LE : pulaar	32	49	17
Nourriture consommée aujourd'hui	83	77	-6
A reçu des cours particuliers en dehors de l'école	76	88	12
Au cours de la dernière semaine complète d'école, combien de jours avez-vous manqué ?	1,55 jour	1,38 jour	-0,17 jour**
Compréhension orale (sur 6)	3,17	3,71	0,54
Conscience phonologique (sur 10)	7,58	7,69	0,11
Connaissance des sons des lettres (sur 100)	75,74	74,98	-0,75
Décodage de pseudo-mots (sur 50)	6,78	7,21	0,43
Décodage de mots familiers (sur 50)	6,16	6,36	0,20
Fluidité de lecture à voix haute (sur 50)	5,37	6,61	1,25
Compréhension écrite (sur 6)	0,56	0,95	0,39
Identifiant numérique (sur 24)	18,19	18,67	0,47

Résultat	Groupe de comparaison	Groupe de traitement	Différence
Comptage (sur 6)	5,12	5,4	0,28
Discrimination des chiffres (sur 12)	9,75	10,37	0,62
Sens des nombres (sur 12)	4,49	5,18	0,69
Addition (sur 12)	6,28	7,17	0,89
Soustraction (sur 12)	4,81	5,24	0,43

Remarque : LE = langue d'enseignement. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D537 . Modèle à effets aléatoires

Résultat	Groupe de comparaison – Situation de référence	Groupe de comparaison – Situation finale	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
Compréhension orale (sur 6)	3,14	4,22	3,40	4,41	-0,06 (0,10)	2 227
Conscience phonologique (sur 10)	7,59	7,35	8,46	8,74	0,15 (0,13)	2 227
Construct de décodage	75,31	77,69	70,75	83,52	0,46*** (0,10)	2 227
• Connaissance des sons des lettres (sur 100)	s/o	s/o	s/o	s/o	0,42*** (0,08)	2 227
• Décodage de pseudo-mots (sur 50)	6,64	9,94	5,99	11,85	0,33*** (0,08)	2 227
• Décodage de mots familiers (sur 50)	6,11	9,00	6,74	12,87	0,48*** (0,10)	2 227
Fluidité de lecture à voix haute (sur 50)	5,24	9,88	5,57	13,12	0,32*** (0,11)	2 227
Compréhension écrite (sur 6)	0,54	0,97	0,68	1,35	0,14* (0,08)	2 191
Numéro d'identification (sur 24)	18,07	20,34	18,23	21,00	0,11 (0,07)	2 227

Comptage (sur 6)	5,11	5,51	5,33	5,69	-0,03 (0,11)	2 227
Discrimination des chiffres (sur 12)	9,74	10,85	10,39	11,22	-0,10 (0,08)	2 227
Sens des nombres (sur 12)	4,43	5,25	4,48	5,84	0,18** (0,08)	2 227
Ajout (sur 12)	6,21	8,27	7,01	9,28	0,06 (0,08)	2 227
Soustraction (sur 12)	4,76	6,22	5,12	7,26	0,20** (0,09)	2 227

Remarque : Les tests portent sur le genre, le niveau scolaire, le fait que l'élève ait bénéficié ou non d'un soutien scolaire externe et la langue dans laquelle l'enquête a été menée. La fluidité de lecture à voix haute était le seul test chronométré, d'une durée de 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D638 . Échantillon complet avec variables de contrôle supplémentaires

Résultat	Groupe de comparaison – Situation de référence	Groupe de comparaison – Situation finale	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
Compréhension orale (sur 6)	3,14	4,22	3,40	4,41	-0,04 (0,10)	2 192
Conscience phonologique (sur 10)	7,59	7,35	8,46	8,74	0,11 (0,13)	2 192
Construct de décodage	s/o	s/o	s/o	s/o	0,41*** (0,08)	2 192
• Connaissance des sons des lettres (sur 100)	75,31	77,69	70,75	83,52	0,41*** (0,09)	2 192
• Décodage de pseudo-mots (sur 50)	6,64	9,94	5,99	11,85	0,34*** (0,09)	2 192
• Décodage de mots familiers (sur 50)	6,11	9,00	6,74	12,87	0,50*** (0,10)	2 192
Fluidité de lecture à voix haute	5,24	9,88	5,57	13,12	0,35*** (0,11)	2 192

Résultat	Groupe de comparaison – Situation de référence	Groupe de comparaison – Situation finale	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
(sur 50)						
Compréhension écrite (sur 6)	0,54	0,97	0,68	1,35	0,16** (0,08)	2 156
Numéro d'identification (sur 24)	18,07	20,34	18,23	21,00	0,09 (0,07)	2 192
Comptage (sur 6)	5,11	5,51	5,33	5,69	-0,01 (0,11)	2 192
Discrimination des chiffres (sur 12)	9,74	10,85	10,39	11,22	-0,11 (0,08)	2 192
Sens des nombres (sur 12)	4,43	5,25	4,48	5,84	0,19** (0,08)	2 192
Ajout (sur 12)	6,21	8,27	7,01	9,28	0,07 (0,09)	2 192
Soustraction (sur 12)	4,76	6,22	5,12	7,26	0,21** (0,09)	2 192

Remarque : Les contrôles comprennent le genre, le niveau scolaire, le fait que l'élève ait bénéficié ou non d'un soutien scolaire externe, la langue dans laquelle l'enquête a été menée, le nombre moyen de jours de cours manqués la semaine précédente déclaré par l'élève, la présence ou non d'un lavabo dans l'école et la présence ou non de savon dans ce lavabo. La fluidité de lecture à voix haute était le seul test chronométré à 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D739 . Échantillon apparié

Résultat	Groupe de comparaison – Situation de référence	Référence finale de comparaison	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
Compréhension orale (sur 6)	3,11	4,22	3,42	4,35	-0,08 (0,11)	1 716
Conscience phonologique (sur 10)	7,53	7,22	8,39	8,79	0,21 (0,13)	1 716
Construct de décodage	s/o	s/o	s/o	s/o	0,41*** (0,08)	1 716

Résultat	Groupe de comparaison – Situation de référence	Référence finale de comparaison	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
• Connaissance des sons des lettres (sur 100)	7	76,48	71,25	83,55	0,52*** (0,10)	1 716
• Décodage de pseudo-mots (sur 50)	6,78	9,7	6,12	11,13	0,27*** (0,09)	1 716
• Décodage de mots familiers (sur 50)	6,29	8,89	6,99	12,51	0,43*** (0,11)	1 716
Fluidité de lecture à voix haute (sur 50)	5,53	9,59	5,64	12,3	0,3** (0,11)	1 716
Compréhension écrite (sur 6)	0,56	0,96	0,72	1,20	0,05 (0,08)	1 694
Numéro d'identification (sur 24)	18,18	20,28	18,36	20,93	0,11* (0,07)	1 716
Comptage (sur 6)	5,09	5,53	5,32	5,71	-0,01 (0,11)	1 716
Discrimination des chiffres (sur 12)	9,71	10,85	10,40	11,19	-0,12 (0,08)	1 716
Sens des nombres (sur 12)	4,43	5,23	4,41	5,61	0,16* (0,09)	1 716
Ajout (sur 12)	6,25	8,20	7,11	9,17	0,06 (0,09)	1 716
Soustraction (sur 12)	4,86	6,25	5,23	7,17	0,16** (0,09)	1 716

Remarque : Les tests portent sur le genre, le niveau scolaire, le fait que l'élève ait bénéficié ou non d'un soutien scolaire externe, et la langue dans laquelle l'enquête a été menée. La fluidité de lecture à voix haute était le seul test chronométré, d'une durée de 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Tableau D840 . Échantillon complet sans les élèves sœurs

Résultat	Groupe de comparaison – Situation de référence	Référence finale de comparaison	Groupe de traitement – Situation de référence	Groupe de traitement – Situation finale	Effet standardisé du traitement	N
Compréhension orale (sur 6)	3,17	4,25	3,71	4,42	-0,21** (0,10)	1 813
Conscience phonologique (sur 10)	7,58	7,4	7,69	8,6	0,31** (0,16)	1 813
Décodage conceptuel	s/o	s/o	s/o	s/o	0,42*** (0,09)	1 813
• Connaissance des sons des lettres (sur 100)	75,74	78,29	74,98	86	0,36*** (0,10)	1 813
• Décodage de pseudo-mots (sur 50)	6,78	10,1	7,21	13,54	0,38*** (0,11)	1 813
• Décodage de mots familiers (sur 50)	6,16	9,04	6,36	12,73	0,50*** (0,13)	1 813
Fluidité de lecture à voix haute (sur 50)	5,37	10,05	6,61	14,63	0,36** (0,15)	1 813
Compréhension écrite (sur 6)	0,56	0,98	0,95	1,56	0,10 (0,10)	1 783
Numéro d'identification (sur 24)	18,19	20,42	18,67	20,92	0,01 (0,08)	1 813
Comptage (sur 6)	5,12	5,51	5,4	5,62	-0,14 (0,11)	1 813
Discrimination des chiffres (sur 12)	9,75	10,85	10,37	11,18	-0,10 (0,09)	1 813
Sens des nombres (sur 12)	4,49	5,32	5,18	6,27	0,10 (0,10)	1 813
Ajout (sur 12)	6,28	8,33	7,17	9,02	-0,05 (0,07)	1 813
Soustraction (sur 12)	4,81	6,26	5,24	6,94	0,06 (0,09)	1 813

Remarque : Les tests portent sur le genre, le niveau scolaire, le fait que l'élève ait bénéficié ou non d'un soutien scolaire externe et la langue dans laquelle l'enquête a été menée. La fluidité de lecture à voix haute était le seul test chronométré, d'une durée de 60 secondes.

* $p < 0,10$. ** $p < 0,05$. *** $p < 0,01$.

Annexe E. Fiabilité inter-évaluateurs

Pour mesurer la fiabilité et le degré d'homogénéité des scores attribués par les enquêteurs lors de l'évaluation des compétences des enfants en **lecture-écriture**, 15 % de l'échantillon de référence et 20 % de l'échantillon de la fin de période ont été évalués simultanément par deux enquêteurs différents. Des techniques d'**analyse de la variance (ANOVA) à un facteur**, de type « long » et à sens unique — utilisées pour déterminer si la moyenne d'une variable dépendante est identique dans deux ou plusieurs groupes non apparentés et indépendants — ont été mobilisées afin de calculer, pour chaque paire d'évaluateurs, la **corrélation intraclasse (ICC)**, comme mesure de la fiabilité inter-évaluateurs. En nous inspirant de Fleiss et al. (1973), nous avons interprété les corrélations intraclasse comme suit :

- Moins de 0,40 – Mauvaise
- Entre 0,40 et 0,75 – Bonne ou passable
- Supérieure à 0,75 – Excellente

Le tableau E1 montre le pourcentage de concordance entre les évaluateurs, ainsi que les scores IRR pour l'échantillon d'évaluation du projet au début et à la fin. Dans l'ensemble, l'IRR pour l'échantillon d'évaluation du projet était excellent pour toutes les variables. Par conséquent, le test ANOVA n'a pas pu calculer l'IRR. Cela signifie que les résultats étaient identiques ; il s'agit d'un résultat positif.

Tableau E141 . Fiabilité inter-évaluateurs par sous-test de compétences en lecture-écriture pour l'échantillon

Sous-test de compétences en lecture-écriture	Référence		Fin	
	IRR	Évaluation	IRR	Note
Âge	s/o	Excellent	s/o	Excellent
Au cours de la dernière semaine complète de cours, combien de jours avez-vous manqué ?	100	Excellent	s/o	Excellent
Nourriture consommée aujourd'hui	96	Excellent	90	Excellent
Fluidité sémantique	97	Excellent	100	Excellent
Conscience phonologique	99	Excellent	100	Excellent
Connaissance des sons des lettres	99	Excellent	100	Excellent
Décodage de pseudo-mots	99	Excellent	100	Excellent
Décodage de mots familiers	100	Excellent	99	Excellent
Fluidité de lecture à voix haute	100	Excellent	100	Excellent
Compréhension écrite	99	Excellent	100	Excellent
Numéro d'identification	s/o	Excellent	s/o	Excellent
Comptage	98	Excellent	90	Excellent

Discrimination des chiffres	s/o	Excellent	s/o	Excellent
Sens des nombres	99	Excellent	s/o	Excellent
Addition	99	Excellent	99	Excellent
Soustraction	98	Excellent	98	Excellent

Remarque : Une valeur de fiabilité IRR indiquée comme s/o signifie qu'aucune variation n'a été observée entre les enquêteurs ; en conséquence, la régression ANOVA n'a pas pu estimer de corrélation infraclass. *N = 165 au début de l'étude et 229 à la fin.*

Annexe F. Instruments d'enquête

Questions préalables à l'enquête

À remplir par l'enquêteur avant l'enquête.

1. Coordonnées GPS (*programmer l'enquête pour collecter les coordonnées GPS*).
2. Sélectionnez la région.
 - a. Diourbel
 - b. Kaolack
 - c. Matam
 - d. Saint-Louis
3. Sélectionnez le nom de l'école. (*Sélectionnez dans la liste déroulante des noms préchargés.*)
4. S'agit-il d'une école de traitement ?
 - a. Oui
 - b. Non

-- Les questions 5 et 6 concernent uniquement l'évaluation finale. --

5. Cet élève a-t-il été interrogé lors de l'évaluation initiale ou s'agit-il d'un élève de remplacement ?
 - a. Élève de l'échantillon de référence
 - b. Remplaçant
6. (*Si la réponse A a été sélectionnée à la question 5.*) Sélectionnez le nom de l'élève. (*Sélectionnez-le dans une liste de noms préchargée. Le programmeur peut filtrer la liste pour n'afficher que les élèves de l'école sélectionnée.*)

--

7. S'agit-il d'une évaluation individuelle ou en binôme ? (À des fins d'assurance qualité)
 - a. Évaluation individuelle ; passer à la section suivante.

- b. Évaluation en binôme
- 8. (Si B est sélectionné pour la question 7) Êtes-vous l'enquêteur chargé de l'observation ou l'enquêteur chargé de l'entretien ?
 - a. Chargé de l'observation
 - b. Chargé de l'entretien

Informations à caractère général

SANS LIMITE DE TEMPS

SANS SUPPORT DE STIMULATION

Lisez chaque question exactement telle qu'elle est écrite. Si l'enfant semble ne pas comprendre, vous pouvez reformuler légèrement la question pour l'aider à comprendre. Laissez également à l'enfant suffisamment de temps pour répondre à chaque question avant de passer à la suivante.

1. Quel est votre nom ? *[Remarque à l'intention de l'enquêteur : veuillez séparer le nom en prénom et nom de famille, le cas échéant. S'il n'y a pas de prénom ou de nom de famille, veuillez inscrire « S/O » dans ce champ].*
2. Sexe de l'enfant :
 - a. Garçon
 - b. Fille
 - c. Autre/inconnu
3. Quel âge avez-vous ? Entrez l'âge en années... *(Écrivez « 999 » si l'âge est inconnu)*
4. En quelle classe êtes-vous actuellement ?
 - a. CP
 - b. CE1
5. Avez-vous suivi un programme d'apprentissage (par exemple, une école maternelle) avant d'entrer au CP ?
 - a. Non
 - b. Oui

- c. Je ne sais pas/sans réponse.
6. Quelle(s) langue(s) parlez-vous à la maison ? (Sélectionnez toutes les réponses applicables.) :
- a. Français
 - b. wolof
 - c. sérère
 - d. pulaar
 - e. Autres, précisez.
7. Quelle(s) langue(s) parlez-vous dans la cour de récréation avec vos amis ? (Sélectionnez toutes les réponses applicables) :
- a. Français
 - b. Wolof
 - c. Sérère
 - d. Pulaar
 - e. Autre, précisez.
8. Au cours de la dernière semaine complète de cours, combien de jours avez-vous manqué ?
- a. 0 jour, j'étais à l'école tous les jours.
 - b. 1 jour
 - c. 2 jours
 - d. 3 jours
 - e. 4 jours
 - f. 5 jours, j'ai manqué tous les jours de cours.
 - g. Je ne sais pas/Pas de réponse.

9. Quelle était la ou les raisons pour lesquelles vous avez manqué l'école ? *(Ne posez cette question que si l'enfant a manqué des jours de cours. Ne lisez pas les choix de réponse. Sélectionnez toutes les réponses applicables.)*

- a. L'enfant était malade.
- b. Pluie / mauvais temps
- c. L'enseignant était absent.
- d. Je ne voulais pas y aller
- e. Je ne pouvais pas payer les fournitures scolaires ou les frais de scolarité.
- f. Je devais aider dans les tâches ménagères
- g. Devais aider mes parents dans leur travail rémunéré/leur emploi ou travaillais pour gagner ma vie.
- h. Problèmes de transport ou absence de moyen de transport.
- i. Je devais m'occuper de mes frères et sœurs
- j. Personne ne m'aide pour que j'aille à l'école.
- k. Autre (précisez)

10. Avez-vous mangé aujourd'hui ?

- a. Oui
- b. Non

11. *(situation de référence uniquement)* Avez-vous participé au programme *Ndaw Wune* après l'école au cours de la dernière année scolaire ? *(Programmeur : posez cette question uniquement aux élèves de CE1 dans les écoles qui ont bénéficié du programme.)*

- a. Oui
- b. Non
- c. Je ne sais pas/Pas de réponse.

(fin de période uniquement) En quel mois avez-vous commencé à suivre les cours *Ndaw Wune* ? *(Programmeur : posez cette question uniquement aux élèves des écoles participant au programme.)*

- a. Décembre
- b. Janvier
- c. Février
- d. Mars
- e. Avril
- f. Mai
- g. Aucune de ces réponses
- h. Je ne sais pas/Pas de réponse.

11b. À quelle fréquence avez-vous assisté aux cours *Ndaw Wune* ?

- a. Généralement trois jours par semaine
- b. En général deux jours par semaine
- c. En général un jour par semaine
- d. Rarement (moins d'un jour par semaine)

11c. Depuis le début du programme en décembre, combien de cours de *Ndaw Wune* avez-vous manqués ?

- a. Je n'ai jamais manqué de cours.
- b. J'ai manqué 1 à 5 cours.
- c. J'ai manqué 6 à 10 cours.
- d. J'ai manqué plus de 10 cours.

12. *(Situation de référence)* Au cours de l'année écoulée, avez-vous suivi des cours particuliers en dehors de l'école ? *(Enquêteur : vous pouvez expliquer que le terme « cours particuliers » désigne un cours qu'un enseignant, un membre de la famille, un ami ou un autre élève qui donne des cours supplémentaires en dehors des heures de classe, en lecture-écriture ou mathématiques. Ces cours s'ajoutent aux cours particuliers Ndaw Wune si l'élève en a bénéficié.)*

(fin de période) Au cours de l'année scolaire actuelle, avez-vous reçu des cours particuliers en dehors des heures de classe, en plus de *Ndaw Wune* (le cas échéant) ?

(Enquêteur : vous pouvez expliquer que le terme « cours particuliers » désigne un cours qu'un enseignant, un membre de la famille, un ami ou un autre élève qui donne des cours supplémentaires en dehors des heures de classe, en lecture-écriture ou mathématiques).

- a. Oui
- b. Non
- c. Je ne sais pas/Pas de réponse

12b. (si l'élève répond oui) Combien d'heures de cours particuliers avez-vous reçues au cours de l'année écoulée ?

- a. Moins de 5 heures au total
- b. 5 à 10 heures au total
- c. Plus de 10 heures au total
- d. Je ne sais pas/Pas de réponse

Évaluation de la lecture

Compréhension orale

Pour cette activité, les enfants doivent être évalués dans la langue d'enseignement utilisée à l'école (wolof, pulaar ou sérère).

SANS LIMITE DE TEMPS

SANS SUPPORT DE STIMULATION

Je vais maintenant te lire une histoire deux fois. Ensuite, je te poserai quelques questions au sujet de cette histoire. Tu écouteras attentivement, puis tu répondras aux questions du mieux que tu pourras. D'accord ? Tu comprends ce que je te demande de faire ? C'est parti ! Écoute attentivement :

- C'est la fin de l'année.
- Le directeur de l'école prépare la fête avec les enseignants et les élèves.
- Les filles font les courses. Les garçons apportent du bois et de l'eau.
- Les enseignants préparent le repas.
- Nous mangeons, nous dansons, nous jouons. Tout le monde est heureux.

Évaluateur : Répétez l'histoire, puis posez les questions.

1. À quelle période de l'année a lieu la fête ? Correct (*fin de l'année*)/Incorrect/Pas de réponse
2. Que fait le directeur ? Correct (*il prépare la fête*)/Incorrect/Pas de réponse
3. Que font les filles ? Correct (*elles font les courses*)/Incorrect/Pas de réponse
4. Que font les garçons ? Correct (*ils apportent du bois et de l'eau*)/Incorrect/Pas de réponse
5. Qui prépare le repas ? Correct (*les enseignants*)/Incorrect/Pas de réponse
6. Pourquoi tout le monde est-il heureux ? Correct (parce qu'ils chantent, dansent et jouent ; parce que c'est amusant, c'est une fête, etc.)/Incorrect/Pas de réponse]

Fluidité sémantique

REMARQUE : La présente section concerne l'ensemble des élèves.

INSTRUCTIONS À L'ENQUÊTEUR :

La présente section doit être répétée dans chaque langue indiquée par l'enfant aux questions 6 et 7. Testez également l'enfant dans la langue nationale utilisée à l'école (wolof, sérère ou pulaar) si l'enfant n'a pas indiqué ces langues.

Montrez à l'enfant l'image qui comprend une scène avec plusieurs images et demandez-lui d'essayer de dire autant de mots que possible sur la scène - ou tout autre mot qui lui vient à l'esprit - en 60 secondes.

Il est important de noter qu'il ne s'agit PAS d'un test sur les mots figurant dans l'image (c'est-à-dire un test de vocabulaire), mais plutôt d'un test sur le nombre et la vitesse avec lesquels les mots viennent à l'esprit de l'enfant dans une langue donnée. Tant que les mots sont corrects dans la langue dans laquelle le test est effectué, il est bon de les marquer comme corrects. Par exemple, s'il y a un cochon dans l'image et que l'enfant dit « cochon » ou « chien » ou « ferme » ou tout autre mot réel dans cette langue, il doit tout de même être marqué comme correct. Notez que l'enfant n'est pas tenu de nommer tous les objets visibles dans la scène.

Notez le nombre de mots prononcés par l'élève pendant les 60 secondes. Après une pause, le test est répété avec le même enfant, en utilisant la même scène mais dans une autre langue, et un nouveau score est enregistré.

NOTES POUR LE PROGRAMMEUR : Nous avons 2 images que nous devons recycler en fonction du nombre de langues. Par exemple, si l'enfant dit qu'il/elle parle :

- 1 langue, nous utiliserons l'image 1.

- 2 langues, nous utiliserons l'image 1 pour la première langue et l'image 2 pour la deuxième langue.
- 3 langues, nous utiliserons l'image 1 pour la première langue, l'image 2 pour la deuxième langue et à nouveau l'image 1 pour la troisième langue.
- 4 langues, nous utiliserons alors l'image 1 pour la première langue, l'image 2 pour la deuxième langue, à nouveau l'image 1 pour la troisième langue, puis à nouveau l'image 2 pour la quatrième langue
- Et ainsi de suite...

Image 1



Image 2



Conscience phonologique

REMARQUE : La présente concerne l'ensemble des élèves.

Pour cette activité, les enfants doivent être testés dans la langue d'enseignement utilisée à l'école (wolof, pulaar ou sérère).

SANS LIMITE DE TEMPS

SANS SUPPORT DE STIMULATION

RÈGLE D'ARRÊT : Si l'élève ne donne aucune réponse correcte parmi les cinq premiers mots, demandez-lui gentiment d'arrêter et cochez la case « Arrêt ». Passez à l'exercice suivant.

ADMINISTRATION : L'enquêteur ne doit cliquer sur l'option de la tablette que lorsque l'enfant ne donne pas de réponse correcte. Si l'enfant donne la bonne réponse, ne cliquez pas sur l'option de la tablette.

[INSTRUCTION À L'ATTENTION DU PROGRAMMEUR : INCLURE L'ARRÊT AUTOMATIQUE SI LES CINQ PREMIÈRES OPTIONS SONT INCORRECTES]

Cet exercice **se fait uniquement à l'oral. Ne montrez pas les mots écrits à l'élève.** Le test n'est pas chronométré. Donnez d'abord à l'élève les exemples les uns après les autres en lui expliquant la tâche. Après chaque exemple, demandez-lui de répondre par lui-même.

INSTRUCTIONS À L'ÉLÈVE :

Cet exercice est un exercice oral. Je vais vous dire un mot deux fois, puis je veux que vous me disiez le tout premier son du mot que vous entendez, d'accord ?

Par exemple : le mot « soupe » commence par le son « ssss », n'est-ce pas ? Quel est le tout premier son du mot « soupe » ? « Soupe » ? [Attendez que l'élève répète le son « ssss ». S'il ne répond pas, dites : « Le tout premier son du mot « soupe » est « ssss »].

Essayons quelques autres exemples :

Quel est le tout premier son du mot « jour » ? Si l'élève répond correctement, dites « Très bien ! Le premier son du mot « jour » est « jjjj ». Si l'élève ne répond pas, dites « Le premier son du mot « jour » est « jjjj ».

Quel est le tout premier son du mot « chic » ? Si l'élève répond correctement, dites « Très bien ! Le premier son du mot « chic » est « ch ». Si l'élève ne répond pas, dites « Le premier son du mot « chic » est « ch ».

Quel est le tout premier son du mot « poulet » ? Si l'élève répond correctement, dites « Très bien ! Le premier son du mot « poulet » est « p ». Si l'élève ne répond pas, dites « Le premier son du mot poulet est « p ».

Comprenez-vous ce que je vous demande de faire ? Je vais maintenant vous lire d'autres mots. Je lirai chaque mot deux fois. Écoutez attentivement et dites-moi le tout premier son que vous entendez, d'accord ?

Ne corrigez pas l'élève pendant le test. S'il ne répond pas ou hésite, après 3 secondes, cochez la case « Pas de réponse » et passez à l'élément suivant.

Règle d'arrêt : si l'élève ne donne aucune réponse correcte parmi les cinq premiers mots, demandez-lui gentiment d'arrêter et cochez la case « Arrêt ». Passez à l'exercice suivant.

1. jour	/j/
2. dur	/d'/
3. lac	/l/
4. sac	/s/
5. car	/k'/
6. par	/p'/
7. tour	/t'/
8. fil	/f/
9. balle	/b'/
10. vol	/v/

Connaissance des lettres

REMARQUE : La présente concerne l'ensemble des élèves.

Pour cette activité, les enfants doivent être testés dans la langue d'enseignement utilisée à l'école (wolof, pulaar ou sérère).

LE SUPPORT DE STIMULATION DOIT ÊTRE FOURNI À L'ÉLÈVE.

RÈGLE D'ARRÊT : Arrêtez le test si l'élève commet 10 erreurs consécutives.

ADMINISTRATION :

Suivez l'écran et marquez toute lettre incorrecte en appuyant dessus à l'écran : elle deviendra bleue. Marquez les corrections automatiques comme correctes en appuyant à nouveau sur la lettre : elle redeviendra blanche.

Restez silencieux, sauf si l'enfant hésite pendant 3 secondes. Si l'enfant hésite pendant 3 secondes, pointez la lettre suivante et dites « **Continue, s'il te plaît** ». Marquez la lettre ignorée comme incorrecte.

Si l'élève prononce le nom de la lettre au lieu de son, dites-lui : « **S'il te plaît, dis-moi le son de la lettre** ». Cet avertissement ne peut être donné qu'une seule fois pendant l'exercice. Si le temps imparti est écoulé avant que l'élève ait lu le dernier élément, l'écran clignotera en rouge et le chronomètre s'arrêtera. Dites à l'enfant d'arrêter.

Notez le nombre total de lettres lues et le nombre de lettres mal lues.

INSTRUCTIONS À L'ÉLÈVE :

Voici une page remplie de lettres de l'alphabet. Dis-moi le SON d'autant de lettres que tu connais, pas leur NOM, mais le SON qu'elles produisent. Par exemple (pointez la lettre C), le son de cette lettre est parfois /k/ comme dans concombre ou parfois /s/ comme dans citron.

Essayons ensemble ! Dis-moi le son de cette lettre (montre la lettre V). Si l'enfant répond correctement, dis : Bien ! Le son de cette lettre est /v/. Si l'enfant ne répond pas correctement, dis : le son de cette lettre est /v/.

Essayons une autre lettre ! Dis-moi le son de cette lettre (montre la lettre L). Si l'enfant répond correctement, dis : « Bien ! Le son de cette lettre est /l/. Si l'enfant ne répond pas correctement, dis : « Le son de cette lettre est /l/.

Tu as compris ?

Quand je dirai « commence », tu vas me dire le son des lettres, aussi vite que possible, mais en faisant attention. Tu vas commencer à dire le son de la lettre en commençant ici et en continuant ainsi. (Indiquez la première lettre de la première ligne et passez votre doigt dans le sens de la lecture de la première ligne.)

Si tu tombes sur une lettre que tu ne connais pas, passe à la lettre suivante. Pose ton doigt sur la première lettre. Es-tu prêt ? Commence !

E	f	O	i	ou	A	z	é	ch	Q
b	on	s	N	T	i	G	m	L	an
w	g	ou	O	M	L	p	T	j	c
V	a	R	K	b	u	s	f	é	J
s	c	an	L	e	D	a	Y	f	H
i	U	p	s	P	M	n	v	oi	T
Z	e	g	un	v	in	an	F	d	o
d	b	A	é	r	m	o	on	T	C
R	q	B	L	ou	e	p	n	i	a
gn	ch	V	E	x	d	m	U	ç	oi

Décodage

REMARQUE : La présente concerne l'ensemble des élèves.

Pour cette activité, les élèves doivent être testés dans la langue d’enseignement utilisée à l’école (wolof, pulaar ou sérère).

Lecture de mots familiers

LE SUPPORT DE STIMULATION DOIT ÊTRE FOURNI À L’ÉLÈVE.

RÈGLE D’ARRÊT : Si l’élève ne donne aucune réponse correcte à cinq questions consécutives, demandez-lui gentiment d’arrêter et cochez la case « Arrêt ». Passez à l’exercice suivant.

ADMINISTRATION :

Suivez l’écran et marquez tout mot incorrect en appuyant dessus à l’écran : il deviendra bleu. Marquez les corrections automatiques comme correctes en appuyant à nouveau sur le mot : il redeviendra blanc.

Restez silencieux, sauf si l’enfant hésite pendant 5 secondes. Si l’enfant hésite pendant 5 secondes, pointez le mot suivant et dites « **Continue, s’il te plaît** ». Marquez le mot ignoré comme incorrect.

[INSTRUCTIONS DU PROGRAMMEUR : INCLUEZ LES AUTRES INSTRUCTIONS CONFORMÉMENT AUX NOTES D’ADMINISTRATION CI-DESSUS. NOTEZ LE NOMBRE TOTAL DE MOTS LUS, LE NOMBRE DE MOTS LUS INCORRECTEMENT ET LE TEMPS TOTAL NÉCESSAIRE POUR LIRE LES MOTS.]

Voici une page avec des mots que vous allez lire. Voici 3 exemples. Le premier mot [indiquez le mot « ta » avec le doigt] se lit « ta ». Peux-tu me lire ce premier mot ? [Après la réponse, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] Et ce mot ? [indiquez le mot « par » avec votre doigt]. Peux-tu me lire ce mot ? [Après la réponse, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] Et ce mot ? [Indiquez le mot « lune » avec votre doigt]. Peux-tu me lire ce premier mot ? [Après la réponse, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] D’accord ?

Tu comprends ce que je te demande de faire ? Quand je dirai « Commence », tu liras les mots de gauche à droite, ligne par ligne. À la fin d’une ligne, tu passeras à la ligne suivante. Essaie de lire rapidement et correctement. Tu es prêt ? Commence.

Démarrez le chronomètre lorsque l’élève essaie le premier mot (« vous ») en appuyant sur le bouton.

tu	ou	il	ma	vol
sa	ami	or	car	lire
sol	sage	peur	bébé	papa
carte	blé	cri	fleur	vache
sur	vole	chaise	bleu	peau
mil	clé	mur	monde	table
fin	posé	date	kilo	tour
ronde	faire	pré	porter	abri
été	rougir	beau	moto	pain
mal	vélo	douze	vide	bol

Lecture de pseudo-mots

LE SUPPORT DE STIMULATION DOIT ÊTRE FOURNI À L'ÉLÈVE.

RÈGLE D'ARRÊT : Si l'élève ne donne pas une seule réponse correcte pour cinq éléments consécutifs, demandez-lui gentiment d'arrêter et cochez la case « Arrêt ». Passez à l'exercice suivant.

ADMINISTRATION :

Suivez les instructions à l'écran et marquez tout mot incorrect en appuyant dessus à l'écran : il deviendra bleu. Marquez les corrections automatiques comme correctes en appuyant à nouveau sur le mot : il redeviendra blanc.

Restez silencieux à moins que l'enfant n'hésite pendant 5 secondes. Si l'enfant hésite pendant 5 secondes, pointez le mot suivant et dites « **Continue, s'il te vous plaît** ». Marquez le mot ignoré comme incorrect.

[INSTRUCTIONS AU PROGRAMMEUR : INCLURE L'AUTOSTOP SI LES CINQ PREMIÈRES OPTIONS SONT INCORRECTES. INCLURE LES AUTRES INSTRUCTIONS CONFORMÉMENT AUX NOTES D'ADMINISTRATION CI-DESSUS. NOTER LE NOMBRE TOTAL DE MOTS LUS ET LE NOMBRE DE MOTS LUS INCORRECTEMENT.

Voici quelques mots que tu n’as peut-être jamais vus. Mais j’aimerais que tu essayes de les lire. Par exemple, ce premier mot [indiquez le mot « bi » avec votre doigt] se lit « bi ». Peux-tu lire ce premier mot ? [Après la réponse, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] Et ce mot ? [indiquez le mot « sep » avec votre doigt]. Peux-tu me lire ce premier mot ? [Après qu’il/elle ait répondu, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] Et ce mot ? [Indiquez le mot « lar » avec votre doigt]. Peux-tu me lire ce premier mot ? [Après la réponse, ou après 5 secondes s’il n’y a pas de réponse, montrez-lui comment faire.] D’accord ?

Tu comprends ce que je te demande de faire ? Quand je dirai « Commence », tu liras les mots de gauche à droite, ligne par ligne. À la fin d’une ligne, tu passeras à la ligne suivante. Essaie de lire rapidement et correctement. Tu es prêt ? Commence.

Démarrez le chronomètre lorsque l’élève essaie le premier mot (« ja ») en appuyant sur le bouton.

ja	tal	zi	ol	vaf
sar	cla	vor	ciko	ul
bige	plovi	neul	bilba	ima
tipa	blu	osi	toche	flir
saré	rané	nur	pro	duse
mouli	clo	chane	doupé	bape
til	zopé	taindé	nube	doul
donré	raite	dreu	lorpe	ibrau
oti	nogir	neau	moudir	bir
bair	lépa	zode	fipe	nour

Fluidité de lecture à voix haute

REMARQUE : La présente section est destinée aux élèves qui ont obtenu une note supérieure à 5 dans la section précédente.

Pour cette activité, les élèves doivent être testés dans la langue d’enseignement utilisée à l’école (wolof, pulaar ou sérère).

DURÉE : 60 secondes

LE SUPPORT DE STIMULATION /CARTE D’AFFICHAGE DOIT ÊTRE FOURNI(E) À L’ÉLÈVE.

RÈGLE D’ARRÊT : Si l’élève ne lit pas un seul mot parmi les mots requis pour répondre à la première question, demandez-lui gentiment d’arrêter et cochez la case « Arrêt ». Passez à l’exercice suivant.

ADMINISTRATION : Démarrez le chronomètre lorsque l’enfant lit le premier mot.

Suivez l’écran et marquez tout mot incorrect en appuyant dessus à l’écran : il deviendra bleu. Marquez les corrections automatiques comme correctes en appuyant à nouveau sur le mot : il redeviendra blanc.

Restez silencieux, sauf si l’enfant hésite pendant 5 secondes. Si l’enfant hésite pendant 5 secondes, pointez le mot suivant et dites « **Continue, s’il te plaît** ». Marquez le mot ignoré comme incorrect.

Si le temps est écoulé avant que l’élève ait lu le dernier mot, l’écran clignotera en rouge et le chronomètre s’arrêtera. Dites à l’enfant d’arrêter. Vérifiez que le dernier mot lu s’affiche à l’écran afin qu’un crochet rouge apparaisse. Appuyez ensuite sur « **Suivant** ».

Si l’enfant atteint le dernier élément avant que l’écran ne devienne rouge/le chronomètre ne s’arrête, arrêtez le chronomètre dès que l’enfant lit le dernier mot. Appuyez sur le dernier mot pour faire apparaître le crochet rouge. Appuyez ensuite sur « **Suivant** ».

[INSTRUCTIONS AU PROGRAMMEUR : INCLURE UN ARRÊT AUTOMATIQUE SI TOUS LES MOTS NÉCESSAIRES POUR RÉPONDRE À LA PREMIÈRE QUESTION SONT MAL LUS. INCLURE LES AUTRES INSTRUCTIONS CONFORMÉMENT AUX NOTES D’ADMINISTRATION CI-DESSUS. NOTER LE NOMBRE TOTAL DE MOTS LUS, LE NOMBRE DE MOTS MAL LUS ET LE TEMPS PRIS.

METTRE EN ŒUVRE LA LOGIQUE SUIVANTE POUR QUE LES QUESTIONS DE LA SECTION SUIVANTE S’AFFICHENT :

- Si l’enfant n’est capable de lire que la première ligne (jusqu’à ou avant « s’amuse »), n’affichez que la première question, à savoir « Quel jour s’amusent-ils ?
- Si l’enfant est capable de lire uniquement les deux premières lignes (jusqu’à ou avant « frère »), affichez uniquement les deux premières questions.
- Si l’enfant est capable de lire uniquement les trois premières lignes (jusqu’à ou avant « un ballon »), n’affichez que les trois premières questions.
- Si l’enfant est capable de lire uniquement les quatre premières lignes (jusqu’à ou avant « au ballon »), ne montrez que les quatre premières questions.

- Si l'enfant est capable de lire les cinq premières lignes (jusqu'à ou avant « contenus »), montrez toutes les questions.

Maintenant, j'aimerais que tu lises à haute voix l'histoire d'un enfant. Essaie de lire rapidement et correctement ; ensuite, je te poserai quelques questions. Commence quand je te le dirai. (Place la carte devant l'élève. Pointe la première ligne (le titre). Es-tu prêt ?

Commencez. [Démarrez le chronomètre en appuyant sur le bouton START / STOP]

Le jeu de ballon / C'est samedi. Ali s'amuse	8
dans la cour de l'école avec son petit frère.	17
Son camarade Sidi arrive avec un ballon.	24
Ali, son frère et Sidi jouent au ballon	32
toute la matinée. Les amis ont bien joué. Ils sont contents.	43

Compréhension écrite

REMARQUE : La présente section s'adresse aux élèves qui ont lu à une vitesse d'au moins 10 mots par minute dans la section précédente.

Pour cette activité, les élèves doivent être testés dans la langue d'enseignement utilisée à l'école (wolof, pulaar ou sérère).

SANS LIMITE DE TEMPS

STIMULUS DE L'ACTIVITÉ PRÉCÉDENTE

[INSTRUCTIONS AU PROGRAMMEUR : LES QUESTIONS DU PROGRAMME DOIVENT APPARAÎTRE EN FONCTION DE LA PARTIE DU PASSAGE LU À HAUTE VOIX QUE L'ENFANT A LU DANS LA SECTION PRÉCÉDENTE.]

Lorsque l'élève a terminé la lecture de la section ci-dessus, retirez-lui le texte et posez-lui la première question ci-dessous. Si l'élève ne répond pas dans les 10 secondes, répétez la question et accordez-lui 5 secondes supplémentaires pour répondre. S'il ne répond toujours pas, passez à la question suivante. Posez-lui des questions qui correspondent aux lignes de texte jusqu'à la ligne où se trouve le crochet (]), c'est-à-dire jusqu'à l'endroit où l'élève a arrêté de lire.

Notez les réponses de l'élève dans l'espace « Réponses de l'élève » : cochez la case qui correspond à sa réponse pour chaque question.

- « Correct » : l'élève donne une réponse correcte ou a donné une réponse incorrecte mais s'est corrigé par la suite (autocorrection). Les réponses correctes peuvent être fournies en français ou dans la langue nationale. Les réponses dont le sens est similaire à celles indiquées doivent être considérées comme correctes.
- « Incorrect » : l'élève donne une réponse incorrecte ou dit « je ne sais pas ».
- « Pas de réponse » : l'élève ne donne pas de réponse.

Vous allez maintenant répondre à quelques questions sur l'histoire.

- Quel jour s'amuse-t-il ? [Samedi] Correct/Incorrect/Pas de réponse
- Où sont Ali et son petit frère ? [au tribunal ; à l'école, etc.] Correct/Incorrect/Pas de réponse
- Qu'est-ce que Sidi apporte avec lui ? [un ballon] Correct/Incorrect/Pas de réponse
- Qui joue au ballon ? [Ali ; Sidi ; son frère ; Ali et Sidi ; etc.] Correct/Incorrect/Pas de réponse
- Comment sont les enfants à la fin du jeu ? [heureux] Correct/Incorrect/Pas de réponse
- Pourquoi sont-ils heureux ? [Ils ont bien joué ; etc.] Correct/Incorrect/Pas de réponse

Évaluation mathématique

Pour l'évaluation mathématique, les enquêteurs doivent expliquer les activités en utilisant le langage le plus familier à l'enfant, selon les résultats de la fluidité sémantique. Les réponses dans n'importe quelle langue sont acceptables.

Identification des nombres

RÉSULTAT D'APPRENTISSAGE Identifie les nombres jusqu'à 999

MATÉRIEL REQUIS

Grille numérique



Script de l'enquêteur sur le terrain : Faisons une activité avec les chiffres.

Voici quelques nombres. Je veux que vous pointiez chaque nombre et que vous me disiez de quel nombre il s'agit. Êtes-vous prêt ? D'accord. Commencez ici.

Question A : Quel est ce chiffre ?

(Remarque : si un enfant s'arrête sur un chiffre pendant 15 secondes, pointez du doigt le chiffre suivant et demandez-lui : « Quel est ce chiffre ? »)

Règles d'arrêt : si l'enfant donne moins de 4 réponses correctes entre les lignes 1 et 2, attribuez-lui un 0 et passez à l'activité suivante.

Si l'enfant passe aux lignes 3 et 4, mais donne moins de 4 réponses correctes entre ces lignes, attribuez-lui un 1 et passez à l'activité suivante.

Ligne 1	6	8	4	2
Ligne 2	1	9	7	10
Ligne 3	1	1	1	20
	2	4	9	
Ligne 4	2	3	6	99
	6	2	0	
Ligne 5	1	4	9	106
	0	0	0	
	0	0	0	
Ligne 6	3	2	8	686
	4	1	3	
	7	4	9	

Réponse attendue : L'enfant lit à haute voix le nombre selon la convention de dénomination des nombres dans la langue d'enseignement ou la langue maternelle.

Instructions pour l'évaluation

Programmeur / Enquêteur : Veuillez noter le nombre de chiffres de chaque ligne auxquels l'élève répond correctement. Indiquez la ou les langues utilisées par l'enfant pour identifier les chiffres.

Langue(s) de réponse : _____

Grille d'évaluation

Lignes 1-2 : Capacité à identifier les nombres jusqu'à 10

Lignes 3-4 : Capacité à identifier les nombres jusqu'à 99

Lignes 5-6 : Capacité à identifier les nombres jusqu'à 999

Comptage

RÉSULTAT

D'APPRENTISSAGE

Compte les nombres jusqu'à 15

MATÉRIEL NÉCESSAIRE

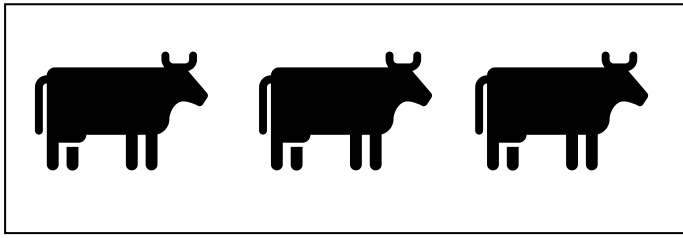
Images



Script de l'enquêteur sur le terrain : Passons à une autre activité. Je vais te montrer quelques images et je voudrais que tu comptes le nombre d'éléments que tu vois dans chacune d'elles.

Voici un exemple.

Combien de vaches vois-tu sur l'image ?



(Si l'enfant répond 3) Bravo, il y a trois vaches sur la photo. Essayons avec d'autres photos.

(Si l'enfant donne un autre nombre ou ne répond pas, pointez chaque vache et comptez.)
Il y a trois vaches sur cette image. Essayons-en d'autres.

(Remarque : si l'enfant reste bloqué sur un problème pendant 30 secondes, passez au suivant).

Images

Image 1

Combien de girafes vois-tu ici ?

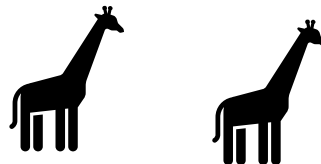


Image 2

Combien d'oiseaux vois-tu ici ?

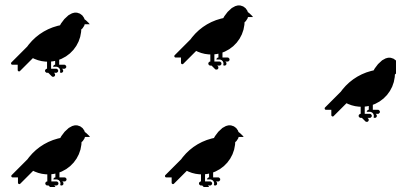


Image 3

Combien de crayons
vois-tu ici ?

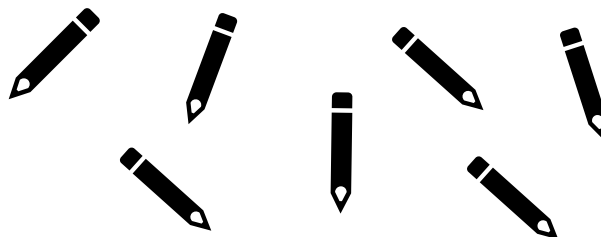


Image 4

Combien de voitures
vois-tu ici ?

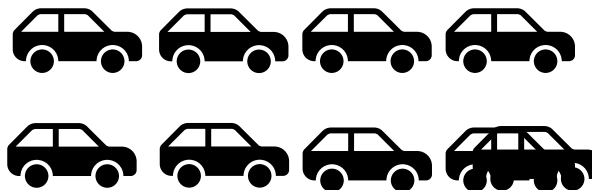


Image 5

Combien d'arbres
vois-tu ici ?

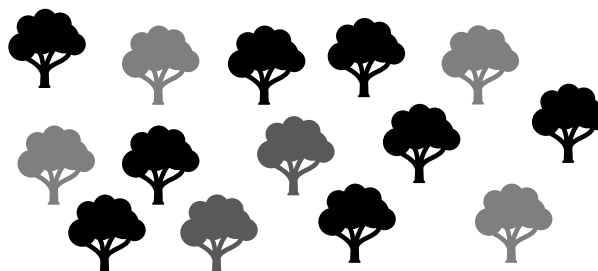
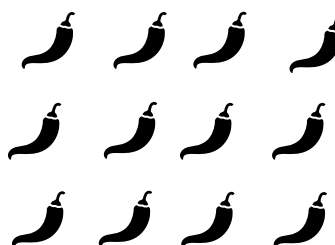


Image 6

Combien de poivrons
vois-tu ici ?



Réponses attendues :

Image 1	2
Image 2	5
Image 3	7
Image 4	9
Image 5	14
Image 6	12

Instructions pour l'évaluation

Programmeur/énumérateur : Notez le nombre de réponses correctes.

Grille d'évaluation

Images 1-2 : Capacité à compter des quantités jusqu'à 5

Images 3-4 : Capacité à compter des quantités jusqu'à 10

Images 5-6 : Capacité à compter des quantités jusqu'à 15

Discrimination des nombres

RÉSULTAT D'APPRENTISSAGE AGE	Compare les nombres jusqu'à 999
MATÉRIEL MATÉRIEL	Grille de discrimination des nombres



Script de l'enquêteur sur le terrain : Faisons une autre activité. Dans cette activité également, je vais te montrer quelques chiffres.

Voici un exemple :

Je veux que tu me dises quel nombre est le plus grand. **(Écrivez les nombres sur une feuille et montrez-les à l'enfant)**

6 4

(si l'enfant pointe le 6) C'est exact, 6 est plus grand. Continuons.

(si l'enfant pointe le 4 ou ne répond pas ; pointez le 6) 6 est plus grand. Essayons encore.

Maintenant, regarde ces chiffres. Dis-moi lequel est le plus grand. **(Écris les chiffres sur une feuille et montre-les à l'enfant)**

13 19

(si l'enfant pointe le 19) C'est exact, 19 est plus grand.

(si l'enfant désigne le 13 ou ne répond pas, désignez le 19) 19 est plus grand.

Maintenant, tu vas essayer tout seul. Regarde ces chiffres. Dis-moi lequel est le plus grand.
(Montrez les chiffres à l'enfant. Pour les chiffres restants, n'indiquez PAS si l'enfant a raison ou tort et ne donnez pas la réponse. Contentez-vous de prendre note de sa réponse et passez à la question suivante.)

(Remarque : si un enfant s'arrête sur une paire de chiffres pendant 15 secondes, pointez la combinaison de chiffres suivante et demandez-lui quel chiffre est le plus grand.

Règles d'arrêt : *si l'enfant donne moins de 2 réponses correctes entre les lignes 1 et 2, attribuez-lui un 0 et passez à l'activité suivante.*

Si l'enfant passe aux lignes 3 et 4, mais donne moins de 2 réponses correctes entre ces lignes, attribuez-lui un 1 et passez à l'activité suivante.

À propos de l’American Institutes for Research®

Fondé en 1946, l’American Institutes for Research® (AIR®) est un organisme à but non lucratif, indépendant et non partisan, spécialisé dans la recherche en sciences comportementales et sociales et fournit une assistance technique aux États-Unis et à l’étranger dans les domaines de l’éducation, de la santé et de la main-d’œuvre. Les interventions d’AIR s’inscrivent dans le cadre de sa mission qui consiste à générer et à utiliser des données rigoureuses contribuant à un monde meilleur et plus équitable. Basée à Arlington, en Virginie, AIR dispose de bureaux aux États-Unis et à l’étranger. Pour plus d’informations, visitez le site [AIR.ORG](https://www.air.org).



Siège social d’AIR®

1400 Crystal
Arlington,
+1.202.403.5000 | [AIR.ORG](https://www.air.org)

Drive,
VA

10e

étage
22202-3289

Avis concernant les marques commerciales : « American Institutes for Research » et « AIR » sont des marques déposées. Tous les autres noms de marques, de produits ou de sociétés sont des marques commerciales ou des marques déposées de leurs propriétaires respectifs.

Copyright © 2025 American Institutes for Research®. Tous droits réservés. Aucune partie de cette publication ne peut être reproduite, distribuée ou transmise sous quelque forme ou par quelque moyen que ce soit, y compris par photocopie, enregistrement, affichage sur un site web ou tout autre moyen électronique ou mécanique, sans l’autorisation écrite préalable de l’American Institutes for Research. Pour toute demande d’autorisation, veuillez utiliser le formulaire « Contactez-nous » sur [AIR.ORG](https://www.air.org).