

Cadre de référence des tests PASEC2024

de lecture et de mathématiques de fin de collège



pasec
Programme d'analyse des systèmes
éducatifs de la confemen

Cadre de référence des tests PASEC2024

de lecture et de mathématiques de fin de collège

Équipe du PASEC

Le présent document a été préparé par l'équipe du Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC).

- Hilaire Hounkpodoté, Coordonnateur du PASEC
- Bassile Tankeu, Chef de la division « Statistiques, analyse des données et études » du PASEC
- Labass Lamine Diallo, Chef de la division « Conception des instruments, collecte des données et renforcement des capacités » du PASEC
- Ousmane Birba, Conseiller technique du PASEC
- Seydou Garba Hamidou, Conseiller technique du PASEC
- Tchindebé Berkeunbé, Conseiller technique du PASEC
- Valérie Vinuesa, Conseillère technique du PASEC
- Fatimé Zara Boukar, Conseillère technique du PASEC
- Adama Bologo, Chargé de projet du PASEC
- Khady Mbaye Camara, Assistante administrative principale du PASEC
- Penda Seck Fall Diouf, Assistante administrative adjointe du PASEC
- Ulrich Tadajeu Kenfack, Attaché de communication du PASEC

Relecture

Alexis LOYE, Président du comité scientifique du PASEC

Publié en 2026 par le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC)

Complexe Sicap Point E, Immeuble C, 3^e étage, Avenue Cheikh Anta Diop, BP : 3220 Dakar, Sénégal

© PASEC 2026

Ce document est disponible en accès ouvert sous licence Creative Commons Attribution – Partage dans les mêmes conditions 4.0 International – CC BY-SA 4.0, sauf indication contraire.

Toute utilisation, reproduction, adaptation, traduction ou diffusion, totale ou partielle, est autorisée à condition de mentionner la source, de citer correctement la publication et de diffuser toute œuvre dérivée sous une licence équivalente.

Pour citer ce document :

PASEC. 2026. *Cadre de référence des tests PASEC2024 de lecture et de mathématiques de fin de collège*. Dakar : PASEC.

La CONFEMEN encourage l'utilisation et la diffusion des informations contenues dans cette publication à des fins de recherche, de formation, d'élaboration de politiques publiques et de plaidoyer en faveur d'une éducation de qualité.

Conception graphique et mise en page : Suzanne Grossmann

ISBN : 92-9133-216-X

Sommaire

Liste des tableaux	4
Sigles et acronymes	4
Introduction	5
Chapitre 1. Cadre de l'évaluation en lecture	7
1.1. Types de textes, supports et domaines évalués en lecture	8
1.1.1. Compréhension des textes narratifs	9
1.1.2. Compréhension des textes argumentatifs	10
1.1.3. Compréhension des textes descriptifs	10
1.2. Processus cognitifs évalués en lecture	11
1.2.1. Compréhension explicite	11
1.2.2. Inférence simple	12
1.2.3. Inférence complexe	12
1.2.4. Interprétation – combinaison – résumé	13
1.3. Exemples d'items de l'évaluation PASEC2024 en lecture – compréhension de texte	13
1.3.1. Exemples d'items pour les différents processus cognitifs	13
1.3.2. Analyse des facteurs cognitifs contribuant au niveau de difficulté de chaque type de question	15
1.4. La présentation des niveaux de compétence des élèves en lecture dans l'évaluation PASEC2024	15
Chapitre 2. Cadre de l'évaluation en mathématiques	17
2.1. Présentation des domaines d'évaluation en mathématiques	18
2.1.1. Arithmétique et Algèbre	19
2.1.2. Mesure	19
2.1.3. Géométrie	19
2.1.4. Statistiques	19
2.2. Processus cognitifs évalués en mathématiques	20
2.2.1. Connaître et comprendre	20
2.2.2. Appliquer des procédures	20
2.2.3. Résoudre des problèmes	20
2.3. Exemples d'items d'évaluation en mathématiques	21
2.3.1. Exemples concrets mettant en œuvre les trois processus cognitifs (connaître des concepts ; appliquer des procédures ; résoudre des problèmes)	22
2.3.2. Analyse des facteurs cognitifs contribuant au niveau de difficulté de chaque type de question	23
2.4. Présentation des niveaux de compétence en mathématiques dans les évaluations PASEC	24

Chapitre 3. Processus de construction des instruments et de collecte des données	27
3.1. Conception des tests	28
3.1.1. Approche participative dans la conception des tests	28
3.1.2. Accompagnement technique et renforcement des capacités	29
3.1.3. Processus de révision et mise à l'essai des items	29
3.2. Validation des instruments	29
3.2.1. Validation qualitative : le laboratoire cognitif	29
3.2.2. Ajustements et révisions des items	30
3.2.3. Validation quantitative : pré-test et analyses psychométriques	30
3.3. Structure et montage final des épreuves PASEC2024 : méthode des cahiers tournants	30
3.4. Administration des tests PASEC2024	31
Chapitre 4. Pistes d'amélioration et perspectives	33
4.1. Réflexions sur l'évolution des cadres d'évaluation	34
4.1.1. L'évolution vers des évaluations plus complètes et diversifiées	35
4.1.2. Intégration des compétences transversales et socio-émotionnelles	35
4.1.3. Évolutions liées aux technologies numériques	35
4.2. Propositions pour l'évolution du cadre d'évaluation du PASEC	36
4.2.1. Renforcer l'évaluation de la résolution de problèmes complexes	36
4.2.2. Diversifier les formats d'items pour stimuler différentes formes de pensée	36
4.2.3. Intégrer l'évaluation des compétences transversales et socio-émotionnelles	37
4.2.4. Utiliser les technologies numériques pour évaluer des processus cognitifs dynamiques	37
Conclusion	38
Bibliographie générale	39

Liste des tableaux

TABEAU 1	Poids des domaines de lecture dans le test PASEC2024	9	TABEAU 5	Pourcentages et nombres d'items des processus cognitifs en mathématiques	20
TABEAU 2	Poids des processus cognitifs et des domaines en lecture dans l'évaluation PASEC2024	11	TABEAU 6	Pourcentages et nombres d'items des processus cognitif dans le test de mathématiques de fin de collège de l'évaluation PASEC2024.....	21
TABEAU 3	Échelle de compétences PASEC2024 en lecture – Fin de collège	16	TABEAU 7	Échelle de compétences PASEC2024 en mathématiques – Fin de collège	24
TABEAU 4	Pourcentages et nombres d'items dans le test de mathématiques de fin de collège	19			

Sigles et acronymes

AERA	American Educational Research Association	OCDE	Organisation de Coopération et de Développement Économique
APA	American Psychological Association	ODD	Objectifs de Développement Durable
APC	Approche par les compétences	PASEC	Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN
BIBD	Balanced Incomplete Block Design	PIRLS	Progress in International Reading Literacy Study
CONFEMEN	Conférence des Ministres de l'Éducation des États et Gouvernements de la Francophonie	PISA	Programme international pour le suivi des acquis des élèves
GAML	Global Alliance to Monitor Learning	QCM	Questionnaire à choix multiple
IEA	International Association for the Evaluation of Educational Achievement	SOLO	Structure of Observed Learning Outcomes
IRT	Item Response Theory	TIMSS	Trends in International Mathematics and Science Study, an international assessment of student achievement in mathematics and science
ITC	International Test Commission	UNESCO	United Nations Educational, Scientific and Cultural Organization
LMTF	Learning Metrics Task Force	UNICEF	Fonds des Nations Unies pour l'enfance
MMS	Multiple Matrix Sampling		
NCME	National Council on Measurement in Education		
NCTM	National Council of Teachers of Mathematics		
NRC	National Research Council		

Introduction

L'amélioration de la qualité des apprentissages constitue aujourd'hui l'un des principaux défis des systèmes éducatifs à l'échelle mondiale. Si les progrès réalisés au cours des dernières décennies ont permis d'accroître considérablement l'accès à l'éducation, ils ne se sont pas systématiquement traduits par une amélioration des acquis scolaires. Les travaux de la Banque mondiale, de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), du Fonds des Nations Unies pour l'enfance (UNICEF) et de l'Organisation de coopération et de développement économiques (OCDE) mettent en évidence une crise des apprentissages qui touche particulièrement les pays à revenu faible ou intermédiaire. Cette crise se traduit par l'incapacité d'une proportion importante d'élèves à acquérir les compétences fondamentales attendues en lecture et en mathématiques, pourtant indispensables à la poursuite des apprentissages et à l'insertion sociale et professionnelle (Banque mondiale et al., 2022 ; OCDE, 2023 ; UNESCO, 2024).

Dans ce contexte, les évaluations internationales à grande échelle jouent un rôle essentiel dans le suivi de la qualité des systèmes éducatifs. Elles permettent de mesurer les acquis des élèves selon des cadres conceptuels communs, de produire des données comparables entre pays et dans le temps, ainsi que d'identifier les facteurs individuels, scolaires et contextuels associés aux performances des élèves. Ces informations constituent un appui important pour le pilotage des politiques éducatives et le suivi de l'Objectif de développement durable 4 (ODD 4), qui vise à assurer une éducation inclusive, équitable et de qualité et à promouvoir les possibilités d'apprentissage tout au long de la vie pour tous (Organisation des Nations Unies, 2015 ; UNESCO, 2024).

En Afrique subsaharienne francophone, le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC), créé en 1991, constitue le principal dispositif régional d'évaluation comparative des apprentissages. Depuis plus de trois décennies, il accompagne les pays membres de la CONFEMEN dans la production de données fiables et comparables sur les acquis scolaires des élèves, tout en contribuant à l'analyse de l'efficacité, de l'équité et

de la qualité des systèmes éducatifs. Les résultats des différentes évaluations PASEC alimentent les réflexions nationales sur les politiques éducatives et soutiennent la mise en œuvre de réformes fondées sur des données probantes (CONFEMEN, 2020).

L'édition PASEC2024 marque une nouvelle étape dans le développement du programme en intégrant une évaluation des élèves en fin de collège. Cette étape du parcours scolaire revêt une importance particulière puisqu'elle correspond à l'achèvement du premier cycle de l'enseignement secondaire. Les compétences acquises à ce niveau conditionnent largement la poursuite des études dans le second cycle du secondaire, l'accès à la formation professionnelle ou l'entrée progressive dans le monde du travail. Elles constituent également un indicateur de la capacité des systèmes éducatifs à préparer les jeunes aux exigences académiques, professionnelles et citoyennes du XXI^e siècle.

L'évaluation porte sur deux domaines disciplinaires fondamentaux : la lecture et les mathématiques. En lecture, elle vise à apprécier la capacité des élèves à comprendre, interpréter et mobiliser des informations issues de différents types de textes. En mathématiques, elle évalue les compétences développées dans les principaux domaines enseignés au collège, notamment l'arithmétique et l'algèbre, la géométrie, les mesures et les statistiques, en mobilisant différents processus cognitifs tels que la connaissance, l'application et la résolution de problèmes.

Le présent cadre de référence expose les fondements conceptuels, curriculaires et méthodologiques qui sous-tendent l'évaluation PASEC2024 en fin de collège. Il précise les domaines de compétences évalués, les processus cognitifs retenus, les principes ayant guidé l'élaboration des instruments de mesure ainsi que les modalités de construction des épreuves. Il constitue ainsi un document de référence destiné à assurer la validité, la fiabilité et la comparabilité internationale des résultats, tout en facilitant leur interprétation et leur utilisation dans le pilotage des politiques éducatives.

Chapitre 1

Cadre de l'évaluation en lecture

Les modèles cognitifs contemporains, notamment le modèle de construction-intégration de Kintsch (1988), définissent la lecture comme un processus actif de construction de sens où le lecteur établit des connexions entre les informations explicites et implicites du texte et ses connaissances antérieures.

L'évaluation des compétences en lecture représente un levier fondamental dans l'analyse des acquis scolaires, en particulier dans les systèmes éducatifs d'Afrique subsaharienne, où l'enjeu de la maîtrise des apprentissages fondamentaux demeure central. Dans le cadre du PASEC2024, l'évaluation de la lecture en fin de collège repose exclusivement sur la mesure du domaine de la compréhension de l'écrit. Ce choix méthodologique vise à garantir la comparabilité entre pays, en s'affranchissant des spécificités linguistiques propres à chaque système éducatif.

La conception de cette évaluation repose sur un double ancrage. D'une part, elle s'appuie sur les travaux scientifiques en matière de lecture et de compréhension de texte et d'autre part, tient compte des contextes sociolinguistiques et culturels des pays participants, qui conditionnent fortement la manière dont les élèves s'approprient la langue d'enseignement et interagissent avec les textes.

Dans un contexte international où la notion de littératie s'impose de plus en plus comme un standard pour juger du niveau de compétence en lecture, le PASEC inscrit son évaluation dans une dynamique comparable. Ainsi, les orientations proposées par la *Learning Metrics Task Force* (LMTF) en matière de mesure de la lecture, et élargies par l'Alliance mondiale pour le suivi des apprentissages (GAML) pour l'ensemble du parcours scolaire – ont servi de repère pour la construction du cadre PASEC2024. Toutefois, l'évaluation PASEC va au-delà d'un simple alignement sur les standards internationaux en développant sa propre échelle de mesure, conçue de manière à refléter de façon fiable les niveaux de compétence atteints par les élèves en compréhension de l'écrit, tout en intégrant la réalité de la transition linguistique et scolaire à la fin du premier cycle du secondaire.

Cette échelle tient donc compte :

- de la maîtrise progressive de la langue d'enseignement par les élèves, dans des contextes où elle n'est souvent pas la langue maternelle ;
- de la diversité des expériences scolaires et sociales qui influencent la capacité des élèves à comprendre, interpréter et mobiliser des textes ;
- et des attentes internationales pour le suivi de l'indicateur 4.1.1 des Objectifs de développement durable (ODD), tout en y adaptant une perspective africaine contextualisée.

L'évaluation s'appuie sur une sélection rigoureuse de textes représentatifs des genres scolaires attendus au collège : textes narratifs, argumentatifs, descriptifs, ainsi que des textes illustrés ou informatifs, comme des fiches ou notices. Ces textes sont choisis pour leur potentiel à mobiliser des processus cognitifs variés, allant de la compréhension littérale à l'inférence simple ou complexe, en passant par l'interprétation et la combinaison d'idées.

Le processus de sélection des textes et de conception des items est lui-même ancré dans une démarche participative et contextualisée. Il mobilise les équipes techniques nationales des pays participants, qui travaillent en collaboration avec des experts du PASEC et des spécialistes en psychométrie et en didactique de la lecture. Tous les items en lecture reposent sur la compréhension de textes de longueurs variables, qui sont choisis avec soin pour éviter les biais culturels, sociaux ou religieux, en veillant à leur accessibilité pour des publics scolaires d'horizons variés. Cette collaboration prend une dimension nouvelle dans le cadre du PASEC2024, qui intègre pour la première fois des pays anglophones et lusophones, nécessitant une attention accrue à la transversalité des références culturelles et à la cohérence des épreuves traduites et adaptées.

Le cadre de l'évaluation en lecture permet donc de prendre en compte la complexité des compétences de compréhension en fin de collège, tout en offrant une mesure rigoureuse, pertinente et sensible aux diversités linguistiques et pédagogiques qui caractérisent les systèmes éducatifs africains. Il reflète une volonté claire de positionner les évaluations du PASEC non seulement comme un outil de comparaison, mais aussi comme un instrument de pilotage éducatif ancré dans les réalités du terrain et conforme aux standards internationaux.

1.1. Types de textes, supports et domaines évalués en lecture

Le cadre d'évaluation du PASEC2024 en lecture est structuré autour de trois grands types de textes : les **textes narratifs**, les **textes argumentatifs** et les **textes descriptifs** (voir *Tableau 1*). Cette organisation reflète la diversité des supports écrits auxquels les élèves sont confrontés dans les programmes scolaires ainsi que dans leur vie quotidienne. Chacun de ces types de textes présente des caractéristiques discursives spécifiques et mobilise des processus cognitifs particuliers, permettant ainsi d'évaluer un large éventail de compétences en compréhension de l'écrit (Duke & Cartwright, 2021).

Le recours à plusieurs types de textes répond à une double exigence. D'une part, il permet de couvrir les principales situations de lecture rencontrées au collège ; d'autre part, il contribue à apprécier la capacité des élèves à adapter leurs stratégies de compréhension en fonction de la structure, de l'organisation et de la finalité du texte. Les questions de compréhension, principalement présentées sous forme de questions à choix multiple, sont élaborées de manière à évaluer différents processus cognitifs, allant du repérage d'informations explicites à l'interprétation, à l'inférence et à l'évaluation critique des informations contenues dans les textes.

Les supports retenus sont représentatifs des écrits auxquels les élèves sont habituellement exposés au cours de leur scolarité. Ils comprennent principalement des textes de longueur moyenne ou importante appartenant aux genres narratifs, argumentatif et descriptif. Cette diversité permet d'assurer une couverture équilibrée des compétences en lecture attendues à la fin du collège tout en tenant compte des exigences des curricula des pays participants.

TABLEAU 1 : Poids des domaines de lecture dans le test PASEC2024

Domaine de lecture	Sous domaines/ Supports de lecture	Proportion dans le test
Compréhension de textes	Compréhension de textes narratifs	50 %
	Compréhension de textes descriptifs	34 %
	Compréhension de textes argumentatifs	16 %

Une attention particulière est portée au choix des thèmes abordés. Les textes sont sélectionnés de manière à être culturellement pertinents et accessibles aux élèves des différents pays participant au PASEC. Ils s’inspirent de situations issues de la vie scolaire, sociale ou quotidienne et sont élaborés de façon à limiter les biais culturels susceptibles d’influencer les performances des élèves. Le processus de développement des instruments associe étroitement les pays participants afin de garantir la pertinence linguistique, culturelle et curriculaire des supports proposés.

Conformément aux recommandations internationales relatives à l’évaluation des compétences en lecture, la classification des supports en textes narratifs, argumentatifs et descriptifs permet d’assurer une représentation équilibrée des principaux genres textuels étudiés au collège et de mesurer les différentes dimensions de la compréhension de l’écrit (Mullis et al., 2019 ; OCDE, 2023). Cette diversité contribue à renforcer la validité de construit de l’évaluation en couvrant les principales situations de lecture auxquelles les élèves sont susceptibles d’être confrontés dans leur parcours scolaire et au-delà.

Enfin, la combinaison de genres textuels variés, de processus cognitifs diversifiés et de formats d’items standardisés permet au PASEC2024 de produire une mesure fiable, valide et comparable des compétences en lecture des élèves de fin de collège, conformément aux standards internationaux en matière d’évaluation à grande échelle (OCDE, 2023).

L’évaluation PASEC2024 accorde une plus grande place au domaine de la compréhension des textes narratifs. Ce choix reflète la configuration des enseignements en fin de collège. Ce domaine constitue une modalité fondamentale de la lecture scolaire et de la vie quotidienne des élèves. Les récits sont omniprésents dans les manuels scolaires, les œuvres littéraires et la culture générale des élèves.

Les textes descriptifs occupent une place plus marginale dans les évaluations de lecture au terme de l’éducation de base, car ils sont souvent considérés comme moins complexes sur le plan cognitif que les textes narratifs ou argumentatifs. Les textes descriptifs sont principalement centrés sur la représentation d’objets, de lieux, de personnes ou de situations sans qu’il y ait nécessairement une progression logique ou une structure narrative.

Enfin, bien qu’essentiels dans la construction de la pensée critique et dans l’expression d’opinions informées, la compréhension des textes argumentatifs occupe une place moins importante dans l’évaluation PASEC2024 en fin de collège, pour plusieurs raisons : complexité cognitive élevée ; les textes argumentatifs exigent des compétences plus avancées en matière d’analyse critique, de déduction, et d’évaluation des raisonnements. Les élèves doivent non seulement comprendre les arguments, mais aussi analyser les contre-arguments, identifier les biais et évaluer la validité des raisonnements (Newell et al., 2011). Ces compétences, bien que cruciales à long terme, sont souvent mieux développées dans les cycles supérieurs du secondaire, où la pratique des débats, des essais et des arguments écrits devient plus systématique.

Alors que la narration et la description sont des formes courantes dans le programme scolaire des collégiens, les textes argumentatifs restent moins fréquents, surtout en début de collège, où les élèves se concentrent davantage sur l’acquisition des compétences de base en lecture. Leur traitement dans les évaluations en fin de collège sert à initier les élèves à des compétences plus complexes, mais elles ne sont pas encore au cœur des objectifs évaluatifs à ce stade.

La mise en exergue dans le test PASEC des textes narratifs par rapport aux textes descriptifs et argumentatifs reflète aussi l’alignement de l’évaluation avec les compétences transversales essentielles au parcours éducatif des élèves. La capacité à comprendre et à interpréter des récits est liée à des compétences critiques pour l’analyse de textes dans toutes les disciplines, ainsi que pour l’acquisition de connaissances dans des contextes académiques variés. Les textes narratifs, en particulier, permettent de tester des compétences qui se traduisent au-delà de la lecture, telles que :

- la compréhension du sens global d’un texte ;
- l’interprétation des intentions et des émotions des personnages ;
- la capacité à tirer des conclusions ou à anticiper des événements futurs ;
- la construction d’une narration ou d’un raisonnement logique.

1.1.1. Compréhension des textes narratifs

Les textes narratifs occupent une place centrale dans le développement des compétences de compréhension en lecture chez les élèves (Duke & Cartwright, 2021). Ils permettent aux élèves de construire une représentation cohérente d’un récit en mobilisant simultanément des connaissances linguistiques, discursives et cognitives. La compréhension des récits constitue ainsi un levier important du développement de la compréhension de l’écrit et des apprentissages scolaires plus largement (Duke & Cartwright, 2021 ; OCDE, 2023).

La compréhension de textes narratifs mobilise plusieurs processus cognitifs complémentaires. Les élèves doivent être capables :

- d'identifier les principaux éléments constitutifs du récit, notamment les personnages, le contexte, les événements, l'intrigue ainsi que les relations qui les unissent (McNamara & Magliano, 2009) ;
- d'établir les relations de causalité et de temporalité entre les événements afin de construire une représentation cohérente de l'histoire (Kintsch, 1998 ; McNamara & Magliano, 2009) ;
- d'interpréter les intentions, les motivations, les croyances et les émotions des personnages, processus qui contribue au développement des capacités inférentielles et de la compréhension des états mentaux d'autrui (Mar et al., 2021 ; Oatley, 2016) ;
- de produire des inférences permettant de relier les informations explicites aux connaissances antérieures, d'anticiper les développements du récit et d'en construire une compréhension globale (Cain & Oakhill, 2007 ; Duke & Cartwright, 2021).

De plus, la narration constitue un format qui engage les élèves à manipuler une variété de compétences langagières et cognitives, ce qui est essentiel pour leur développement académique (Snow, 2010). En fin de collège, ces compétences sont cruciales, car elles forment la base de la compréhension de textes plus complexes dans des matières telles que les sciences humaines, l'histoire et la géographie.

Selon les recherches en didactique, notamment celles de Giasson (2000), la compréhension des récits est un des premiers leviers pour développer des compétences en inférence et en anticipation, des compétences essentielles à la lecture.

Dans le cadre du PASEC2024, les textes narratifs visent à évaluer la capacité des élèves à construire une représentation cohérente du récit, à comprendre les relations entre les personnages et les événements, à établir des inférences et à interpréter le sens global du texte. Les élèves sont ainsi amenés à mobiliser des processus de compréhension allant du repérage d'informations explicites à l'interprétation des relations implicites qui structurent le récit.

1.1.2. Compréhension des textes argumentatifs

Les textes argumentatifs mobilisent des processus de compréhension plus élaborés que le simple repérage ou l'organisation des informations. Leur compréhension repose sur la capacité des élèves à analyser un raisonnement, à établir des relations logiques entre les idées, à identifier la thèse défendue ainsi que les arguments et les preuves qui la soutiennent. Elle exige également d'évaluer la pertinence, la cohérence et la crédibilité des informations présentées, compétences qui constituent des composantes essentielles de la pensée critique et de la littératie en lecture (Britt & Rouet, 2012 ; Duke & Cartwright, 2021).

La compréhension des textes argumentatifs mobilise notamment les capacités suivantes :

- identifier la thèse ou le point de vue défendu par l'auteur ;
- distinguer les faits, les opinions, les arguments et les exemples utilisés pour soutenir une idée ;
- établir les relations logiques entre les différentes propositions du texte ;
- produire des inférences afin de relier les informations explicites aux connaissances antérieures ;
- évaluer la cohérence, la pertinence et la force des arguments développés.

Ces processus sollicitent des compétences de raisonnement, d'analyse critique et d'évaluation qui dépassent la simple compréhension littérale du texte. Ils contribuent au développement de la capacité des élèves à interpréter des informations complexes, à apprécier la validité d'un raisonnement et à construire un jugement éclairé, compétences reconnues comme essentielles pour la poursuite des études, la citoyenneté et l'apprentissage tout au long de la vie (OCDE, 2023 ; UNESCO, 2024).

Dans le cadre du PASEC2024, les textes argumentatifs permettent d'évaluer la capacité des élèves à comprendre la structure d'une argumentation, à identifier les liens entre les arguments et les preuves mobilisées, à distinguer les faits des opinions et à interpréter la logique générale du raisonnement développé dans le texte. L'évaluation vise ainsi à apprécier leur aptitude à mobiliser des processus de compréhension de haut niveau, indispensables à la lecture critique de documents de plus en plus présents dans les apprentissages scolaires et dans la vie quotidienne.

1.1.3. Compréhension des textes descriptifs

Les textes descriptifs visent principalement à présenter les caractéristiques d'un objet, d'une personne, d'un lieu, d'un phénomène ou d'une situation. Leur compréhension repose sur la capacité des élèves à repérer, organiser et intégrer des informations afin de construire une représentation cohérente des éléments décrits. Ce type de texte est largement présent dans les disciplines scolaires, notamment en sciences, en géographie et dans les disciplines technologiques, où il constitue un support privilégié pour la transmission de connaissances (Duke & Cartwright, 2021 ; OCDE, 2023).

La compréhension des textes descriptifs mobilise plusieurs processus cognitifs complémentaires. Les élèves doivent notamment être capables :

- d'identifier les informations explicites présentées dans le texte ;
- d'organiser les informations selon leurs relations spatiales, temporelles ou conceptuelles ;

- d'établir des liens entre les différentes caractéristiques décrites afin de construire une représentation globale de l'objet ou du phénomène ;
- de mobiliser leurs connaissances antérieures pour interpréter correctement les informations fournies et en dégager le sens général (Kintsch, 1998 ; McNamara & Magliano, 2009).

Contrairement aux textes narratifs, dont l'organisation repose principalement sur une succession d'événements, les textes descriptifs privilégient une organisation des informations fondée sur les propriétés, les attributs ou les relations entre les éléments décrits. Leur compréhension exige ainsi des compétences de catégorisation, de hiérarchisation et d'intégration de l'information, essentielles à l'acquisition des connaissances dans de nombreuses disciplines scolaires (Duke & Cartwright, 2021).

Dans le cadre du PASEC2024, les textes descriptifs permettent d'évaluer la capacité des élèves à extraire des informations explicites, à établir des relations entre les différents éléments présentés et à construire une représentation cohérente du contenu du texte. Lorsque des illustrations, schémas ou autres supports visuels accompagnent le texte, les élèves doivent également être capables d'intégrer les informations provenant de différentes sources afin d'en construire une compréhension globale, compétence désormais reconnue comme une composante importante de la littératie en lecture (OCDE, 2023).

En définitive, le cadre d'évaluation du PASEC2024 appréhende la compréhension en lecture à travers trois grands types de textes : les textes narratifs, les textes argumentatifs et les textes descriptifs. Chacun de ces genres mobilise des processus cognitifs spécifiques tout en contribuant au développement de compétences communes de compréhension, d'interprétation et de raisonnement. Leur complémentarité permet de couvrir un large éventail des situations de lecture rencontrées au collège et de fournir une mesure valide et équilibrée des compétences en lecture des élèves de fin de collège (AERA et al., 2014 ; Duke & Cartwright, 2021 ; OCDE, 2023).

1.2. Processus cognitifs évalués en lecture

L'évaluation des compétences en lecture dans le cadre du PASEC2024 ne se limite pas à la simple mesure des connaissances des élèves, mais s'intéresse également aux processus cognitifs qu'ils mobilisent pour comprendre les textes. Quatre principaux processus cognitifs sont évalués : la compréhension explicite, l'inférence simple, l'inférence complexe, et l'interprétation-combinaison-résumé. Ces processus permettent de mieux comprendre comment les élèves interagissent avec les textes, et comment ils traitent les informations pour construire du sens. Cette approche repose sur des théories pédagogiques et didactiques reconnues, notamment celles de Kintsch (1998) et de Giasson (2000), qui soulignent l'importance des différents niveaux de traitement de l'information pour une lecture approfondie et critique.

TABLEAU 2 : Poids des processus cognitifs et des domaines en lecture dans l'évaluation PASEC2024

Processus cognitifs évalués	Poids des domaines/supports de lecture de texte dans PASEC2024			Total compréhension de texte
	Textes narratifs	Textes descriptifs	Textes argumentatifs	
Compréhension explicite	10 % (9)	10 % (9)	1 % (1)	21 % (19)
Inférence simple	21 % (19)	12 % (11)	6,5 % (6)	40 % (36)
Inférence complexe	10 % (9)	7 % (6)	1 % (1)	18 % (16)
Interpréter et combiner des informations	9 % (8)	6 % (5)	6,5 % (6)	21 % (19)
Total	50 % (45)	35 % (31)	15 % (14)	100 % (90)

Les caractéristiques principales des trois processus cognitifs sont décrites et illustrées ci-dessous.

1.2.1. Compréhension explicite

La compréhension explicite correspond à la capacité des élèves à repérer, sélectionner et restituer des informations directement exprimées dans un texte, sans qu'il soit nécessaire de réaliser des inférences ou d'établir des liens implicites entre les informations. Elle constitue un processus fondamental de la compréhension en lecture, puisqu'elle permet au lecteur de construire une représentation initiale du contenu du texte, sur laquelle s'appuieront les traitements cognitifs plus complexes, tels que les inférences, l'interprétation et l'évaluation critique (Duke & Cartwright, 2021 ;

Giasson, 2000 ; OCDE, 2023). Ce processus cognitif est souvent considéré comme le premier niveau de compréhension en lecture, car il ne requiert pas de traitement complexe des informations. Les recherches de Giasson (2000) montrent que la compréhension explicite est cruciale pour construire une base solide de compréhension générale du texte avant d'aborder des niveaux de traitement plus élevés. La compréhension explicite mobilise principalement des processus d'identification, de localisation et de reconnaissance de l'information. Bien que ce niveau de compréhension soit souvent considéré comme le plus élémentaire, il constitue une condition essentielle au développement des niveaux supérieurs de compréhension (Kintsch, 1998).

Dans le cadre du PASEC2024, ce processus cognitif est évalué à travers des tâches amenant les élèves à :

- repérer des informations explicites, telles que des personnes, des lieux, des dates, des événements ou des caractéristiques directement mentionnés dans le texte ;
- identifier les principaux éléments constitutifs d'un texte, notamment le thème, les personnages, le lieu, le moment ou les événements décrits ;
- reconnaître une information formulée différemment mais strictement équivalente à une information explicitement présente dans le texte ;
- sélectionner, parmi plusieurs propositions, celle qui correspond fidèlement au contenu littéral du texte.

1.2.2. Inférence simple

L'inférence simple correspond à la capacité du lecteur à déduire des informations qui ne sont pas explicitement formulées dans le texte, mais qui peuvent être établies à partir de la mise en relation d'indices textuels proches. Elle permet d'assurer la continuité de la compréhension en reliant différentes informations explicites afin de construire une représentation cohérente du texte (Kintsch, 1998 ; McNamara & Magliano, 2009). Kintsch (1998) souligne que l'inférence simple est une étape clé dans la construction d'une compréhension plus approfondie, car elle permet au lecteur de combler les lacunes entre les informations données dans le texte et celles qui doivent être imaginées ou interprétées.

En fin de collège, l'élève dans sa lecture de texte entame une construction du sens qui le conduit à nuancer davantage son analyse pour réaliser des inférences directes et en déduire des informations complémentaires en se basant sur le sens littéral d'une phrase ou d'un paragraphe. Le jeune lecteur perçoit des éléments concrets qui lui permettent de construire du sens par déduction et raisonnement logique. Il met en relation plusieurs phrases et idées fournies par le texte pour déduire le sens d'une expression, d'une phrase ou, à un niveau plus poussé, l'idée générale d'un passage dans un texte.

Dans le cadre du PASEC2024, les élèves sont amenés à mobiliser ce processus cognitif pour :

- déduire une information implicite à partir de plusieurs informations explicites présentes dans le texte ;
- établir une relation de cause à effet entre des événements ou des idées ;
- interpréter les conséquences d'une action ou d'un raisonnement à partir des indices fournis par le texte ;
- déduire les intentions, les émotions ou les réactions d'un personnage à partir de ses paroles, de ses actions ou du contexte ;
- identifier l'idée principale d'un passage lorsque celle-ci n'est pas formulée explicitement.

Ces tâches permettent d'évaluer la capacité des élèves à construire le sens du texte au-delà de sa compréhension littérale, en mobilisant des inférences directement fondées sur les informations textuelles disponibles.

1.2.3. Inférence complexe

L'inférence complexe correspond à la capacité du lecteur à construire des interprétations qui dépassent les informations directement fournies par le texte. Elle mobilise simultanément les informations explicites, les relations établies entre différentes parties du texte ainsi que les connaissances antérieures du lecteur afin de construire une représentation globale et cohérente du sens. Contrairement aux inférences simples, qui reposent principalement sur des indices textuels proches, les inférences complexes exigent des processus d'intégration, d'interprétation et de raisonnement plus élaborés (Kintsch, 1998 ; McNamara & Magliano, 2009).

La réalisation d'inférences complexes nécessite notamment de mettre en relation des informations dispersées dans le texte, d'interpréter des informations implicites, de mobiliser des connaissances issues de l'expérience personnelle ou scolaire et d'évaluer les intentions de l'auteur. Ces processus contribuent au développement de la compréhension approfondie des textes, de la lecture critique et du raisonnement (Duke & Cartwright, 2021 ; OCDE, 2023 ; Pressley & Afflerbach, 1995).

À la fin du collège, les élèves devraient être capables de mobiliser ces processus pour construire une compréhension globale du texte et produire des interprétations fondées sur l'ensemble des informations disponibles. Ils doivent notamment être en mesure :

- d'établir des liens entre différentes parties d'un texte afin d'en dégager le sens global ;
- d'interpréter des informations implicites nécessitant la mobilisation de connaissances antérieures ;
- d'identifier les intentions de l'auteur ou le message sous-jacent d'un texte ;
- de comprendre des procédés de langage tels que la métaphore, la comparaison, l'ironie ou l'implicite lorsque leur interprétation nécessite une mise en relation d'informations textuelles et de connaissances extérieures ;
- d'anticiper les conséquences d'une situation ou l'évolution d'un récit en s'appuyant sur l'ensemble des indices disponibles.

Lors de l'évaluation PASEC2024, l'inférence complexe est évaluée à travers des tâches nécessitant de comprendre des significations cachées ou de tirer des conclusions globales à partir de plusieurs éléments dispersés dans le texte. Par exemple, les élèves devront établir des connexions logiques entre différents événements d'un récit ou comprendre le but implicite d'un texte argumentatif. Ils peuvent également comprendre une métaphore ou une ironie en faisant le lien avec des connaissances antérieures ou anticiper la suite de l'histoire en s'appuyant sur les indices présents dans le

texte. Ce processus cognitif est particulièrement pertinent pour évaluer la capacité des élèves à aller au-delà du contenu textuel explicite et à formuler des interprétations nuancées.

1.2.4. Interprétation – combinaison – résumé

Le processus d'interprétation, combinaison et résumé constitue la forme la plus complexe de la compréhension en lecture. Il s'agit pour les élèves de combiner plusieurs informations issues du texte, de les interpréter à la lumière de leurs connaissances préalables et de les synthétiser pour produire un résumé cohérent. Ce processus repose sur les théories de la construction du sens en lecture, telles que celles de van Dijk et Kintsch (1983), qui soulignent l'importance de l'élaboration et de la hiérarchisation des informations pour obtenir une compréhension globale et critique d'un texte.

Dans l'évaluation PASEC2024, ce processus sera mis en évidence par des tâches qui demandent aux élèves de combiner des informations provenant de différentes sections d'un texte pour en produire une interprétation globale. Ils devront également démontrer leur capacité à synthétiser les points clés d'un texte narratif, argumentatif ou descriptif, en les organisant de manière claire et concise. Le résumé d'un texte, en particulier, est une compétence essentielle qui permet de vérifier la compréhension globale d'un texte et la capacité des élèves à hiérarchiser les informations.

Ces quatre processus cognitifs évalués en lecture constituent des indicateurs essentiels de la capacité des élèves à comprendre des textes variés de manière active et approfondie. Ils permettent d'évaluer non seulement la capacité à saisir des informations explicites, mais aussi à effectuer des inférences et à interpréter les informations de manière critique. Les théories

pédagogiques et didactiques (Giasson, 2000 ; Kintsch, 1998 ; Pressley & Afflerbach, 1995) qui sous-tendent ces processus soulignent l'importance de développer des compétences de lecture à plusieurs niveaux, afin de former des lecteurs capables de comprendre, d'analyser, et de réagir aux textes qu'ils rencontrent dans leur parcours scolaire et dans la vie quotidienne. Le PASEC2024 s'appuie sur ces approches pour offrir une évaluation complète des compétences de lecture, en mettant l'accent sur des processus cognitifs qui sont au cœur de la réussite scolaire.

1.3. Exemples d'items de l'évaluation PASEC2024 en lecture – compréhension de texte

Dans l'évaluation des compétences en lecture, les items d'évaluation jouent un rôle central pour mesurer la compréhension et l'interprétation des textes par les apprenants. Ces items, sont conçus pour évaluer différentes dimensions de la lecture, allant de la simple reconnaissance d'informations explicites à l'analyse plus complexe d'inférences et d'opinions. Ils varient en fonction des objectifs pédagogiques et des niveaux de compétence visés, tout en s'adaptant aux divers formats d'évaluation. Dans cette section, plusieurs exemples d'items d'évaluation en lecture sont présentés, illustrant leur structure, leurs objectifs et leur utilité dans l'appréciation des compétences en compréhension écrite.

1.3.1. Exemples d'items pour les différents processus cognitifs

Des exemples concrets mettant en œuvre les quatre processus cognitifs sont présentés dans les pages suivantes.

1

2

3

4

Lisez le texte et répondez aux questions qui suivent.

Je pense sincèrement que nous, les parents africains, nous ne devrions pas laisser sans précaution nos jeunes regarder tous les programmes diffusés par la télévision.

Tout d'abord, parce que comme tous les jeunes du monde, les jeunes enfants africains ne font pas de différence entre ce qu'ils voient en images et la réalité. Dès lors, les vedettes des séries télévisées qui ne sont pas toujours – hélas ! – des exemples de bonne conduite, deviennent pour eux des modèles qu'ils ont tôt fait d'imiter. Ceci explique l'apparition sur le continent africain de certains types de comportements, de certaines façons de s'habiller ou même de parler, qui étaient jusque-là inconnues et même considérées comme inacceptables.

Ensuite, la télévision diffuse à longueur de journée des images de violence de plus en plus insoutenables. Les scènes de guerre y côtoient des scènes de lynchage, qui elles-mêmes succèdent à des scènes de tueries. Les jeunes Africains, qui sont exposés à ces images, grandissent ainsi dans un milieu marqué par la violence, ce qui les prédispose à devenir eux-mêmes plus tard brutaux.

Enfin, on ne tardera pas à découvrir que la télévision est également une des causes non négligeables des mauvais résultats scolaires des jeunes enfants en Afrique, en même temps qu'un facteur de désordre au sein des familles.

Richards, Sciences humaines n°165 octobre 2005

EXEMPLE 1 : Item relatif au processus de Compréhension explicite

De quel continent est l'auteur du texte ?

L95

- A. ☐ D'Afrique
- B. ☐ D'Europe
- C. ☐ D'Asie
- D. ☐ D'Amérique

EXEMPLE 2 : Item relatif au processus d'Inférence simple

Que signifie l'expression "les prédispose à" dans la phrase « les jeunes Africains, qui sont exposés à ces images, grandissent ainsi dans un milieu marqué par la violence, ce qui les prédispose à devenir eux-mêmes plus tard brutaux. »

L96

- A. ☐ Les décourage
- B. ☐ Les empêche
- C. ☐ Les prépare
- D. ☐ Les force

EXEMPLE 3 : Item relatif au processus d'Inférence complexe

Que signifie la phrase « [...] nous ne devrions pas laisser sans précaution nos jeunes regarder tous les programmes diffusés par la télévision » ?

L99

- A. ☐ Il faut priver les jeunes des programmes diffusés à la télévision.
- B. ☐ Il faut aider les jeunes à sélectionner les bons programmes diffusés à la télévision.
- C. ☐ Il faut prendre des mesures pour punir les élèves qui suivent les programmes diffusés à la télévision.
- D. ☐ Il faut laisser les jeunes suivre tous les programmes diffusés à la télévision.

EXEMPLE 4 : Item relatif au processus d'interprétation, combinaison et résumé

Quel titre peut porter ce texte ?

L104

- A. ☐ La violence dans nos sociétés.
- B. ☐ Les avantages des programmes télévisés.
- C. ☐ Le désordre dans les familles africaines.
- D. ☐ Les méfaits des programmes télévisés.

1.3.2. Analyse des facteurs cognitifs contribuant au niveau de difficulté de chaque type de question

Le niveau de difficulté d'un item dans un test de lecture est influencé par une multitude de facteurs que les concepteurs peuvent anticiper et moduler lors du développement des épreuves. Comme le soulignent Haladyna et Rodriguez (2013), la validité et la qualité des items dépendent largement des choix délibérés faits par les concepteurs pour calibrer la complexité des tâches selon les objectifs cognitifs visés.

Dans le cadre des évaluations standardisées, il est essentiel de générer une banque d'items de difficulté variée pour couvrir adéquatement l'ensemble des processus cognitifs liés à la compréhension de l'écrit (Kintsch, 1998 ; Afflerbach, Pearson & Paris, 2008). À travers chaque type de texte évalué (narratif, argumentatif, descriptif), plusieurs éléments peuvent influencer la difficulté d'un item, soit isolément, soit en interaction :

- **La nature du texte.** Plus un texte est long, complexe dans sa structure, dense en informations, ou linguistiquement opaque (par son vocabulaire, la familiarité du thème ou sa charge culturelle), plus les tâches qui s'y rapportent sollicitent des compétences avancées. Les textes nécessitant une navigation cognitive non linéaire, comme ceux comportant des digressions ou des ambiguïtés lexicales, induisent généralement un niveau de difficulté supérieur par rapport à des textes courts et simples, proches du lexique et des structures syntaxiques du quotidien scolaire.
- **Le degré d'explicitation de l'information.** Lorsqu'une tâche implique de localiser une information explicite, littérale et concrète, généralement en début de texte, elle mobilise des processus cognitifs de bas niveau, comme l'identification ou la reconnaissance (Bloom, 1956 ; PIRLS Framework, Mullis et al., 2017). À l'inverse, les tâches requérant l'inférence d'informations implicites, l'interprétation ou la reformulation, relèvent de niveaux cognitifs plus élevés et sont perçues comme plus difficiles.
- **Le type d'appariement exigé :** Selon le modèle de la distance textuelle (Snow, 2002), plus il faut combiner des informations

dispersées dans le texte, voire les croiser avec des éléments contextuels externes, plus la tâche s'éloigne d'un appariement direct et devient cognitivement exigeante. Ainsi, une tâche qui exige la synthèse d'éléments issus de plusieurs paragraphes pour en dégager une idée centrale sera nécessairement plus complexe qu'un appariement simple mot à mot.

- **Le type d'interprétation ou de raisonnement requis :** Plus l'élève est amené à mobiliser des connaissances extérieures, à établir des jugements nuancés ou à confronter différents points de vue implicites dans le texte, plus la tâche suppose une opération cognitive de haut niveau. Ces tâches s'inscrivent dans des dimensions métacognitives et critiques de la compréhension.
- **La proximité entre les choix de réponse :** Dans le cas d'items à choix multiples, la difficulté augmente lorsque les distracteurs sont proches de la bonne réponse, partageant des traits sémantiques ou syntaxiques. Ceci rend la discrimination entre les options plus subtile, et donc l'item plus complexe (Haladyna, Downing & Rodriguez, 2002). À l'inverse, si la bonne réponse est nettement distincte des distracteurs, la tâche est perçue comme plus facile.

En outre, des facteurs **contextuels et logistiques** peuvent affecter la difficulté perçue des items indépendamment de leur contenu : les conditions de passation (bruit, stress, gestion du temps), la durée du test, la qualité matérielle (lisibilité, impression), ou encore l'expérience antérieure des élèves avec les formats d'évaluation utilisés. Ces éléments doivent être pris en compte pour garantir l'équité et la validité de l'évaluation.

1.4. La présentation des niveaux de compétence des élèves en lecture dans l'évaluation PASEC2024

Pour faciliter la lecture et l'interprétation des résultats statistiques, les performances des élèves au test de compréhension de l'écrit de fin de collège sont présentées sur une échelle de compétences segmentées en plusieurs niveaux. Cette toute première échelle de compétence PASEC de lecture *en fin de collège* est calibrée sur la base du fonctionnement des items en vue de présenter les résultats.

1

2

3

4

À chaque niveau de l'échelle correspond un ensemble de connaissances et compétences maîtrisées par les élèves qui relèvent de ce niveau en fonction des réponses au test de tous élèves. Chacune des compétences et des connaissances requises à

chaque niveau est présentée à la page suivante dans l'échelle de compétence PASEC2024. Ces informations sont essentielles pour apprécier le degré de maîtrise des élèves et comprendre les principales difficultés qu'ils rencontrent.

TABEAU 3 : Échelle de compétences PASEC2024 en lecture – Fin de collège

Niveaux	Caractéristique du niveau	Description des compétences
Niveau 4	Compréhension experte et interprétation élaborée	L'élève construit une compréhension approfondie de textes de genres variés, y compris narratifs, descriptifs, argumentatifs et explicatifs. Il mobilise simultanément des informations explicites et implicites, établit des relations entre des éléments parfois éloignés dans le texte et réalise des inférences complexes. Il interprète le texte en prenant en compte le point de vue de l'auteur, les intentions de communication, les sentiments des personnages et les procédés utilisés pour construire le sens. Il est capable d'intégrer différentes informations afin de produire une compréhension globale, nuancée et cohérente.
Niveau 3	Compréhension approfondie et interprétation partielle	L'élève comprend des textes de genres variés et mobilise de manière autonome les informations explicites et implicites pour construire le sens du texte. Il réalise des inférences simples de façon régulière et réussit une partie des inférences plus complexes nécessitant de relier plusieurs informations parfois éloignées. Il commence à interpréter le texte en identifiant les intentions, les sentiments ou certaines caractéristiques des personnages, même si ces interprétations demeurent limitées aux situations les plus accessibles.
Seuil « suffisant » de compétences		
Niveau 2	Compréhension des informations explicites et premières inférences	L'élève comprend les informations explicites présentes dans des textes simples appartenant à différents genres et est capable d'établir des liens directs entre des informations proches afin de produire des inférences simples. Il commence à interpréter certains éléments lorsque ceux-ci sont fortement suggérés par le texte. En revanche, les inférences nécessitant de rapprocher plusieurs informations éloignées ou de mobiliser des connaissances implicites demeurent fragiles. Ce niveau correspond globalement aux compétences minimales attendues à la fin du primaire et constitue le premier niveau de maîtrise permettant une lecture autonome de textes simples.
Niveau 1	Compréhension élémentaire de l'information explicite	L'élève repère et restitue des informations explicitement formulées dans des textes courts, simples et familiers, principalement de type descriptif ou narratif. Il comprend le sens global lorsque les informations sont directement accessibles et fortement contextualisées. En revanche, il éprouve des difficultés à établir des liens entre plusieurs informations ou à produire des inférences, même simples. Ses capacités de compréhension demeurent inférieures aux compétences minimales attendues à la fin du primaire.

Cette échelle est présentée dans les rapports internationaux et nationaux PASEC2024 et permet de présenter et de comparer le pourcentage moyen d'élèves de fin de collège qui atteignent un certain niveau ou le seuil de compétences « suffisant » pour tous les pays ; elle peut être analysée par pays ; par zone éducative dans chaque pays ; par variable d'intérêt traitée par le PASEC.

D'autres informations méthodologiques plus poussées concernant la construction des échelles de compétences sont disponibles dans le rapport technique PASEC2024.

Chapitre 2

Cadre de l'évaluation en mathématiques

Les mathématiques jouent un rôle central dans le développement cognitif des élèves et sont souvent considérées comme un indicateur clé de la réussite scolaire. Elles contribuent au développement du raisonnement logique, de la pensée quantitative, de la modélisation et de la résolution de problèmes, tout en fournissant les outils nécessaires pour comprendre et analyser des situations rencontrées dans la vie scolaire, sociale et professionnelle. La maîtrise des compétences mathématiques constitue ainsi un déterminant important de la réussite scolaire, de la poursuite des études et de la participation à la vie économique et citoyenne (*National Research Council*, 2001 ; Organisation pour la Coopération et le Développement économique [OCDE], 2023).

Dans ce contexte, l'évaluation des compétences en mathématiques au sein du PASEC2024 en fin de collège vise à mesurer la capacité des élèves à mobiliser des connaissances mathématiques essentielles dans divers contextes et à travers différents processus cognitifs. Elle ne se limite pas à mesurer l'acquisition de connaissances factuelles ou la maîtrise de procédures de calcul, mais cherche également à évaluer la capacité des élèves à mobiliser leurs connaissances dans des situations variées, à raisonner mathématiquement et à résoudre des problèmes relevant de contextes scolaires ou de la vie courante. Cette conception de la compétence mathématique est aujourd'hui largement partagée par les principaux cadres internationaux d'évaluation et par la recherche en didactique des mathématiques (Kilpatrick et al., 2001 ; *National Council of Teachers of Mathematics* [NCTM], 2014 ; OCDE, 2023).

La définition du cadre d'évaluation du PASEC2024 repose sur deux sources complémentaires. D'une part, il s'appuie sur les travaux de recherche qui montrent que le développement des compétences mathématiques repose sur l'articulation de plusieurs dimensions complémentaires, notamment la compréhension conceptuelle, la maîtrise des procédures, le raisonnement et la résolution de problèmes (Hiebert & Grouws, 2007 ; Kilpatrick et al., 2001 ; NCTM, 2014). D'autre part, il s'appuie sur les études curriculaires conduites par la CONFEMEN afin d'identifier un socle commun de compétences susceptible d'être évalué de manière comparable dans l'ensemble des pays participants.

À cette fin, la CONFEMEN a réalisé deux études complémentaires : une étude exploratoire des curricula des pays membres en vue de l'élaboration d'un référentiel commun de compétences et une analyse comparative des programmes du premier cycle du secondaire ainsi que des principaux référentiels internationaux en mathématiques. Ces travaux montrent que, malgré des différences dans l'organisation des programmes, la progression des apprentissages et la répartition des contenus selon les niveaux scolaires, les curricula présentent une convergence importante quant aux grands domaines mathématiques enseignés au collège.

L'ensemble des pays participants organise ainsi les apprentissages autour de quatre grands domaines communs : l'arithmétique et l'algèbre, la mesure, la géométrie ainsi que les statistiques et les probabilités, même si l'importance relative accordée à chacun de ces domaines varie selon les systèmes éducatifs. Ce constat a conduit le PASEC à retenir ces quatre domaines comme fondement de l'évaluation, garantissant ainsi à la fois la pertinence curriculaire des contenus évalués et la comparabilité internationale des résultats.

Ce choix est également cohérent avec les principaux cadres internationaux d'évaluation des mathématiques, qui organisent eux aussi les contenus autour de domaines comparables tout en mettant l'accent sur la mobilisation des connaissances dans des situations variées (*International Association for the Evaluation of Educational Achievement* [IEA], 2021 ; OCDE, 2023).

Au sein de chacun de ces domaines, les compétences des élèves sont évaluées selon trois processus cognitifs complémentaires :

- **Connaître**, qui renvoie à la mobilisation de connaissances, de définitions, de propriétés, de faits, de concepts et de procédures mathématiques ;
- **Appliquer**, qui consiste à utiliser ces connaissances et ces procédures dans des situations familières nécessitant la mise en œuvre de démarches appropriées ;
- **Raisonner et résoudre des problèmes**, qui mobilise des capacités d'analyse, de justification, de modélisation, d'interprétation et de résolution de situations nouvelles ou complexes.

Cette classification est directement inspirée du cadre d'évaluation de TIMSS, qui distingue les domaines cognitifs *Knowing, Applying et Reasoning* (IEA, 2021). Elle est également cohérente avec le cadre PISA, qui définit la compétence mathématique comme la capacité à formuler, utiliser et interpréter les mathématiques dans une diversité de contextes afin de résoudre des problèmes et de prendre des décisions éclairées (OCDE, 2023).

Le cadre d'évaluation du PASEC2024 permet ainsi d'apprécier non seulement le niveau de maîtrise des connaissances mathématiques des élèves, mais également leur capacité à mobiliser ces connaissances pour raisonner, interpréter des situations, résoudre des problèmes et utiliser les mathématiques dans des contextes variés. Il fournit aux pays participants des informations comparables sur les acquis des élèves tout en identifiant les domaines et les compétences qui méritent une attention particulière dans les politiques d'amélioration de la qualité de l'enseignement des mathématiques.

2.1. Présentation des domaines d'évaluation en mathématiques

Quatre domaines principaux sont évalués : l'arithmétique et l'algèbre, la mesure, la géométrie et les statistiques. Chacun de ces domaines est structuré en fonction de la progression naturelle des connaissances mathématiques et des compétences cognitives que les élèves doivent acquérir au cours de leur parcours scolaire.

Le cadre de référence de l'évaluation PASEC2024 au collège est adapté à ces différentes exigences et couvre exclusivement des questions de mathématiques se rapportant aux domaines évalués.

Les items du test de mathématiques du PASEC2024 sont classés dans ces quatre domaines de contenu en fonction de ses propres caractéristiques. Par projection, le poids des questions dans le test correspond également aux points qui seront attribués

en mathématiques sur l'échelle de résultat pour chacun des items, des sous-domaines et des domaines.

Les items de test du domaine de l'Arithmétique et l'algèbre couvrent une place plus importante dans le test par rapport aux autres domaines de contenu retenus dans le test de mathématiques (voir *Tableau 4*). Cette répartition dans le cadre de référence des tests PASEC de mathématiques reflète les résultats des analyses curriculaires réalisées par la CONFEMEN auprès des pays participants.t.

TABLEAU 4 : Pourcentages et nombres d'items dans le test de mathématiques de fin de collège

Domaines de mathématiques	Nombre d'items	Poids dans le test
Arithmétique et Algèbre	31	39 %
Mesure	20	25 %
Géométrie	20	25 %
Statistique	9	11 %
Total général	80	100 %

Au-delà de l'appartenance à un des quatre domaines de contenu et sous-domaines, les items de test de mathématiques sont également construits et organisés dans les épreuves PASEC pour couvrir plusieurs champs cognitifs. Le croisement des items selon leur appartenance à un des quatre grands domaines de contenu ainsi qu'à un des trois processus cognitifs tel que défini par le PASEC caractérise les principales propriétés de chaque question au sein des épreuves PASEC.

Les caractéristiques des activités cognitives telles qu'abordées et définies dans les épreuves PASEC2024 sont décrites à la partie suivante.

2.1.1. Arithmétique et Algèbre

L'arithmétique et l'algèbre constituent une base importante des connaissances mathématiques, ce qui les rend essentiels dans toute évaluation en mathématiques. Les recherches montrent que ces compétences sont cruciales pour le développement de la pensée abstraite et pour la résolution de problèmes plus complexes à des niveaux supérieurs (Kilpatrick, Swafford & Findell, 2001 ; OCDE, 2023). Les évaluations dans ce domaine incluent des tâches telles que la résolution d'équations, l'identification de relations entre les nombres et les fonctions de base, qui permettent de tester non seulement la connaissance des opérations numériques, mais aussi la capacité des élèves à généraliser ces concepts dans des contextes divers.

2.1.2. Mesure

Le domaine de la mesure englobe la longueur, la surface, le volume et les unités de mesure, compétences indispensables

dans la vie quotidienne et professionnelle. Selon Clements (2004), la compréhension des concepts de mesure est une composante clé du développement des compétences spatiales et de la pensée mathématique pratique. Ce domaine d'évaluation met l'accent sur la capacité des élèves à utiliser correctement des unités de mesure standards et à appliquer des formules mathématiques pour résoudre des problèmes impliquant des quantités mesurables. Il a été démontré que la maîtrise des concepts de mesure est corrélée avec le succès en sciences physiques et en ingénierie (Smith, 2006).

2.1.3. Géométrie

La géométrie constitue un autre domaine fondamental de l'évaluation, car elle engage les élèves dans la compréhension des formes, des propriétés et des transformations géométriques. Van Hiele (1986) a établi un modèle de développement des compétences géométriques selon lequel les élèves passent par plusieurs niveaux de compréhension, allant de la reconnaissance des formes simples à la capacité d'analyser et de prouver des théorèmes géométriques. De nombreuses recherches ont confirmé la pertinence de ce modèle pour comprendre le développement de la pensée géométrique et continuent de l'utiliser comme cadre de référence pour la conception des curricula et des évaluations en mathématiques (Sriraman & English, 2010). L'évaluation des connaissances géométriques vise à tester la capacité des élèves à manipuler des formes, à appliquer des propriétés géométriques et à résoudre des problèmes dans un espace à deux ou trois dimensions. Il convient de relever cependant que pour de nombreux élèves subsahariens ces notions géométriques telles qu'enseignées sont fondées sur une représentation de l'espace et des formes qui leur sont très peu familière. Les performances des élèves dans les évaluations PASEC du primaire en géométrie et même le fonctionnement psychométrique des items de ce domaine témoignent de cette réalité. Ces compétences sont souvent considérées comme essentielles pour le développement de la pensée logique et pour la résolution de problèmes dans d'autres domaines comme la physique ou l'ingénierie (Battista, 2007).

2.1.4. Statistiques

Le domaine des statistiques vise à développer chez les élèves la capacité à recueillir, organiser, représenter, analyser et interpréter des données afin d'éclairer une prise de décision. Dans une société où les informations quantitatives occupent une place croissante, la culture statistique constitue désormais une compétence essentielle pour comprendre des phénomènes sociaux, économiques, scientifiques et environnementaux, ainsi que pour exercer une citoyenneté éclairée (Franklin et al., 2007 ; OCDE, 2023).

Le domaine des statistiques, de plus en plus important dans les programmes scolaires modernes, mais faiblement représenté dans les curricula en Afrique subsaharienne. Il est intégré à l'évaluation pour mesurer la capacité des élèves à analyser des données, à interpréter des graphiques et des diagrammes, et à tirer des conclusions fondées sur des informations quantitatives. Les recherches en éducation statistique soulignent que l'enseignement des statistiques ne doit pas être envisagé comme une simple

application des mathématiques, mais comme le développement d'un mode de raisonnement spécifique fondé sur l'analyse des données, la prise en compte de la variabilité, de l'incertitude et du contexte (Cobb & Moore, 1997 ; Moore, 1997).

L'évaluation en statistiques dans le cadre du PASEC2024 inclut des tâches qui évaluent la compréhension des concepts de base tels que la moyenne, la médiane, les probabilités, et l'interprétation des représentations graphiques. La maîtrise de ces concepts est également cruciale pour l'enseignement supérieur et des carrières dans les sciences de la donnée, un domaine en pleine expansion.

2.2. Processus cognitifs évalués en mathématiques

L'évaluation des processus cognitifs en mathématiques est essentielle pour comprendre non seulement ce que les élèves savent mais, aussi comment ils utilisent et appliquent ces connaissances dans diverses situations. Les processus cognitifs évalués dans le cadre du PASEC2024 s'articulent autour de trois niveaux principaux : connaître, appliquer et résoudre des problèmes (voir *Tableau 5*). Ces processus sont largement soutenus par les théories de l'apprentissage cognitif et les modèles d'acquisition des compétences mathématiques (Bloom, 1956 ; Anderson & Krathwohl, 2001).

TABLEAU 5 : Pourcentages et nombres d'items des processus cognitifs en mathématiques

Processus cognitif	Nombre d'items	Poids dans le test
Connaître et comprendre les concepts	25	31 %
Appliquer des procédures	33	41 %
Résoudre des problèmes	22	28 %
Total général	80	100 %

2.2.1. Connaître et comprendre

Le processus de **connaître** renvoie à l'identification et à la reconnaissance des concepts mathématiques de base, des formules, et des règles. À ce niveau, les élèves doivent montrer leur capacité à rappeler des faits mathématiques et à reconnaître des principes fondamentaux, ce qui constitue la base de l'apprentissage plus complexe. Les recherches de Hiebert et Carpenter (1992) ont montré que la connaissance des concepts mathématiques est un préalable pour réussir à appliquer ces ceux-ci dans des contextes variés. Dans ce processus cognitif, les élèves peuvent être amenés à identifier des équations, nommer des formes géométriques, ou rappeler des définitions mathématiques standards.

La compréhension conceptuelle constitue une composante essentielle des compétences mathématiques. Elle permet aux élèves de donner du sens aux concepts, aux procédures et aux

relations mathématiques, et soutient la mobilisation des connaissances dans les différents processus cognitifs évalués par le PASEC. Ainsi, bien qu'elle ne constitue pas un processus cognitif distinct, elle est mobilisée à des degrés divers dans les processus « Connaître », « Appliquer » et « Résoudre des problèmes » (*National Council of Teachers of Mathematics [NCTM], 2014*).

Les items de tests appartenant au processus « connaître et comprendre des concepts » occupent une place importante dans le test par rapport aux autres domaines de contenu retenus dans le test de mathématiques comme l'illustre le tableau ci-dessus. Cette dominante dans le cadre de référence des tests PASEC de mathématiques reflète essentiellement l'importance accordée aux différents processus cognitifs dans la structuration des enseignements-apprentissages des mathématiques au collège.

2.2.2. Appliquer des procédures

Le processus cognitif « Appliquer des procédures » se concentre sur la capacité des élèves à utiliser leurs connaissances mathématiques dans des contextes pratiques et familiers. Il suppose que les élèves soient capables de sélectionner et d'utiliser de manière appropriée les concepts, les propriétés, les règles et les procédures mathématiques déjà acquis afin de traiter des situations concrètes ou scolaires ne nécessitant pas un raisonnement particulièrement complexe (IEA, 2021).

Les recherches en didactique des mathématiques montrent que cette capacité à mobiliser des connaissances dans des situations contextualisées constitue une étape essentielle du développement de la compétence mathématique. Elle traduit la capacité des élèves à établir des liens entre les concepts appris et les situations auxquelles ils sont confrontés, favorisant ainsi le transfert des apprentissages vers de nouveaux contextes (Lesh & Doerr, 2003 ; NCTM, 2014). Bien que les travaux de Pólya (1945) et de Schoenfeld (1985) portent principalement sur la résolution de problèmes, ils montrent également que la maîtrise et l'utilisation appropriée des procédures constituent un préalable indispensable à la mise en œuvre de stratégies de résolution efficaces.

Dans le cadre du PASEC2024, ce processus est évalué au moyen de tâches qui demandent notamment aux élèves d'effectuer des calculs numériques, d'appliquer des propriétés géométriques, d'utiliser des formules usuelles, de convertir des unités de mesure, de résoudre des équations simples ou encore de mobiliser les notions de proportionnalité dans des situations proches de celles rencontrées dans les apprentissages scolaires ou dans la vie quotidienne. Ces tâches permettent d'apprécier la capacité des élèves à utiliser leurs connaissances de manière pertinente et efficace dans des contextes familiers, avant d'aborder des situations nécessitant un raisonnement plus élaboré ou une démarche de résolution de problèmes plus complexe.

2.2.3. Résoudre des problèmes

Le processus cognitif « Résoudre des problèmes » constitue le niveau le plus élaboré des compétences évaluées en mathématiques dans le cadre du PASEC2024. Il renvoie à la capacité des

élèves à mobiliser et à coordonner leurs connaissances, leurs procédures et leur raisonnement pour résoudre des situations nouvelles ou complexes, qui ne peuvent être traitées par la simple application d’une procédure connue. Les élèves doivent analyser la situation, sélectionner les informations pertinentes, élaborer une stratégie adaptée, mettre en œuvre une démarche de résolution et interpréter les résultats obtenus (IEA, 2021 ; OCDE, 2023).

Dans son modèle classique, Pólya (1945) décrit la résolution de problèmes comme une démarche organisée comprenant quatre grandes étapes : comprendre le problème, élaborer un plan de résolution, mettre ce plan en œuvre et vérifier la pertinence de la solution obtenue. Bien que ces étapes ne soient pas évaluées séparément dans le PASEC, elles constituent un cadre théorique largement reconnu pour comprendre les processus mobilisés lors de la résolution de problèmes.

Les recherches de Schoenfeld (1985) montrent que les élèves qui maîtrisent les compétences de résolution de problèmes tendent à utiliser une variété de stratégies, telles que la décomposition du problème en sous-problèmes ou l’utilisation d’analogies avec des problèmes déjà rencontrés. De plus, le modèle SOLO (*Structure of Observed Learning Outcomes*) de Biggs et Collis (1982) souligne que les compétences de résolution de problèmes progressent à travers plusieurs niveaux de complexité, allant de l’utilisation de stratégies simples à la construction de réponses intégrées qui montrent une maîtrise globale des concepts mathématiques.

L’évaluation dans ce domaine met l’accent sur des tâches qui ne peuvent être résolues simplement par l’application de routines

pré-apprises. Les élèves doivent démontrer leur capacité à conceptualiser un problème, à sélectionner les informations pertinentes, et à formuler une solution adaptée au contexte. Cela inclut la résolution de problèmes impliquant plusieurs étapes, la modélisation mathématique, et l’application de théories mathématiques pour résoudre des questions complexes dans des domaines tels que la géométrie et l’algèbre.

Le processus «Résoudre des problèmes» représente un poids inférieur car il mobilise des compétences plus complexes et exige une compréhension approfondie des concepts et des procédures. Son évaluation nécessite des tâches plus longues et plus difficiles, qui sont abordées progressivement au collège, après la maîtrise des procédures et des concepts de base.

Dans son ensemble, le cadre d’évaluation du PASEC2024 repose sur une progression des processus cognitifs allant de la mobilisation de connaissances fondamentales (Connaître) à leur utilisation dans des situations familières (Appliquer), puis à leur mobilisation intégrée dans des situations complexes (Résoudre des problèmes). Cette organisation, cohérente avec les principaux cadres internationaux d’évaluation des mathématiques (IEA, 2021 ; OCDE, 2023), permet d’apprécier non seulement les connaissances acquises par les élèves, mais également leur capacité à les mobiliser de manière pertinente, flexible et efficace dans une grande diversité de contextes.

Le *Tableau 6* présente la répartition des questions dans les épreuves définitives PASEC2024 selon les trois processus cognitifs et les quatre domaines et six sous-domaines de contenu.

TABLEAU 6 : Pourcentages et nombres d’items des processus cognitif dans le test de mathématiques de fin de collège de l’évaluation PASEC2024

Domaines et sous-domaine de contenu	Connaître et comprendre des concepts	Appliquer des procédures	Résoudre des problèmes	Total	
				nombre	%
Arithmétique et Algèbre	8	12	11	31	39
Mesure	7	9	4	20	25
Géométrie	6	10	4	20	25
Statistique	4	2	3	9	11
Total général	25	33	22	80	100
Pourcentage	31 %	41 %	28 %	-	-

Le croisement de ces processus avec les domaines de contenu permet d’apprécier globalement la structure du test administré aux élèves en mathématiques et la répartition souhaitée des items de test sur l’échelle de résultat. En effet, cette répartition des items de test vise également à estimer la répartition souhaitée du score entre ces différentes catégories puisque que chaque question est corrigée de façon dichotomique.

2.3. Exemples d’items d’évaluation en mathématiques

L’évaluation des compétences en mathématiques, comme celle mise en place dans le cadre du PASEC2024, repose sur l’utilisation d’items soigneusement conçus pour mesurer divers processus cognitifs tels que connaître, appliquer et résoudre des problèmes. Ces items sont structurés pour couvrir les différents domaines mathématiques – arithmétique, algèbre, géométrie, mesure et

statistiques – et permettent de capturer une vision complète des acquis des élèves en fin de cycle secondaire.

La conception des items s'inspire des meilleures pratiques issues de la recherche en évaluation, en s'appuyant sur des travaux qui montrent que la diversité des types de questions permet de mieux comprendre les compétences spécifiques et les approches cognitives des élèves (Kilpatrick, Swafford & Findell, 2001). En variant les formats de questions, les évaluations offrent un aperçu plus nuancé du niveau de compréhension des concepts mathématiques, et permettent de cibler des compétences spécifiques, qu'il s'agisse de simples rappels de faits mathématiques ou de la résolution de problèmes complexes et contextualisés (Schoenfeld, 1985).

Dans cette section, nous présenterons des exemples d'items pour chaque domaine des mathématiques, en illustrant les différents processus cognitifs évalués. Ces exemples permettront de montrer comment les concepts sont testés de manière progressive, de la reconnaissance des connaissances fondamentales à la mobilisation de compétences pour résoudre des problèmes plus élaborés.

2.3.1. Exemples concrets mettant en œuvre les trois processus cognitifs (connaître des concepts ; appliquer des procédures ; résoudre des problèmes)

EXEMPLE 1 : Item relatif au processus Connaître les concepts

Domaine de l'arithmétique-algèbre

Un agriculteur vend x sacs de riz et y sacs de maïs. La masse de chaque sac de riz est de 50 kg, et la masse de chaque sac de maïs est de 30 kg. Laquelle de ces expressions algébriques représente la masse totale des sacs vendus ?

- A. ☐ $50x + 30$
- B. ☐ $50 + 30y$
- C. ☐ $(50 + 30)(x + y)$
- D. ☐ $50x + 30y$ (Réponse correcte)

EXEMPLE 2 : Item relatif au processus Appliquer des procédures

Domaine de la mesure

Un cylindre a une hauteur de 10 cm et un rayon de 3 cm. Quel est son volume ?

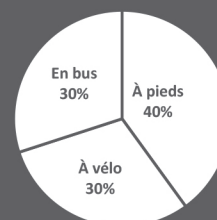
- A. ☐ 90 (Réponse correcte)
- B. ☐ 100
- C. ☐ 80
- D. ☐ 60

EXEMPLE 3 : Item relatif au processus Résoudre des problèmes complexe

Domaine de la statistique

Le diagramme circulaire ci-dessous représente la répartition en pourcentage des moyens par lesquels les 300 élèves d'un petit collège se rendent dans leur établissement.

Répartition des élèves en fonction du moyen de transport



Combien d'élèves utilisent un autre moyen de transport que le bus ?

- A. ☐ 70
- B. ☐ 210 (Réponse correcte)
- C. ☐ 170
- D. ☐ 150

2.3.2. Analyse des facteurs cognitifs contribuant au niveau de difficulté de chaque type de question

Le niveau de difficulté d'un item en mathématiques ne dépend pas uniquement du contenu mathématique mobilisé, mais également des processus cognitifs qu'il sollicite. Dans les évaluations internationales à grande échelle, telles que le PASEC, TIMSS ou PISA, la difficulté des items est construite de manière intentionnelle afin de couvrir un continuum de compétences, depuis la mobilisation de connaissances fondamentales jusqu'au raisonnement et à la résolution de problèmes complexes (IEA, 2021 ; OCDE, 2023). Cette gradation permet de différencier les niveaux de maîtrise des élèves, d'améliorer le pouvoir discriminant des épreuves et d'assurer une mesure fiable des compétences sur l'ensemble de l'échelle d'évaluation.

Les recherches en psychométrie cognitive montrent que la difficulté d'un item résulte davantage des traitements cognitifs exigés que du seul contenu mathématique évalué. Selon Tatsuoka (1983), la résolution d'un item repose sur la mobilisation de plusieurs attributs cognitifs, dont le nombre et la complexité influencent directement sa difficulté. Ainsi, deux items portant sur un même contenu peuvent présenter des niveaux de difficulté très différents selon les connaissances, les procédures et les raisonnements qu'ils requièrent.

Les recherches récentes en didactique et en évaluation des mathématiques confirment que la difficulté d'une tâche dépend également de sa formulation, du contexte dans lequel elle est inscrite, des représentations utilisées (expressions symboliques, graphiques, tableaux, figures) ainsi que du degré de modélisation ou d'interprétation attendu de la part des élèves (Stacey & Turner, 2015 ; OCDE, 2023). Par conséquent, une tâche mobilisant un contenu mathématique relativement simple, comme une addition ou une proportion, peut s'avérer plus difficile qu'un exercice d'algèbre si elle exige un raisonnement non routinier, l'interprétation de plusieurs représentations ou la sélection d'une stratégie de résolution appropriée.

Dans le cadre du PASEC2024, une attention particulière est portée à la progression de la difficulté des items. Celle-ci est construite en faisant varier simultanément le domaine de contenu, le processus cognitif mobilisé, le degré de contextualisation des situations proposées et la complexité des raisonnements attendus. Cette approche permet d'évaluer les compétences mathématiques des élèves sur un large continuum de performance tout en garantissant la validité et le pouvoir discriminant des épreuves.

Dans ce cadre, plusieurs facteurs cognitifs peuvent être identifiés comme déterminants du niveau de difficulté d'un item :

- **Le type de connaissances mobilisées**

Plus un item se limite à la restitution de connaissances factuelles ou de procédures mémorisées (par exemple, une formule ou une règle de calcul simple), plus il est accessible aux élèves. À l'inverse, les items sollicitant une compréhension conceptuelle plus profonde (comme la justification d'une propriété ou l'explication d'un raisonnement) requièrent des compétences de niveau cognitif supérieur (Bloom, 1956 ; Biggs &

Collis, 1982). Ces derniers sont donc généralement perçus comme plus difficiles.

- **Le niveau de familiarité du contexte de la tâche**

Lorsque la situation proposée est proche du vécu quotidien des élèves (par exemple, calculer une distance ou un prix dans une situation familière), la tâche est souvent facilitée. En revanche, un contexte abstrait, ou relevant d'un univers non connu ou non maîtrisé (science, technologie, économie), augmente la charge cognitive, même si les opérations mathématiques requises sont simples (Greer & Verschaffel, 2007). Le décodage de l'énoncé devient alors un obstacle en soi.

- **Le degré de transformation ou de traitement requis**

Un item qui exige une transformation directe d'un nombre ou l'application mécanique d'une opération est généralement plus simple qu'un item qui demande plusieurs étapes de traitement (conversion d'unités, lecture de données, identification de relations, etc.). La complexité cognitive augmente également lorsque plusieurs concepts ou représentations (tableaux, graphiques, figures géométriques) doivent être combinés pour produire une réponse correcte (Leinhardt, Zaslavsky & Stein, 1990).

- **Le raisonnement requis pour résoudre la tâche**

La résolution de problèmes constitue un levier central de différenciation de la difficulté. Plus la tâche nécessite une planification, la mise en œuvre de stratégies ou l'adaptation de procédures dans un contexte nouveau, plus elle implique une activation des processus cognitifs complexes (Schoenfeld, 1985). Ce sont ces types d'items qui permettent d'évaluer en profondeur la capacité de transfert des apprentissages mathématiques.

- **La proximité et la nature des distracteurs (dans les QCM)**

Dans les items à choix multiples, la difficulté est également liée à la qualité des distracteurs. Lorsque ceux-ci sont clairement incorrects ou peu plausibles, la bonne réponse est facilement identifiable. En revanche, des distracteurs proches, qui reprennent des erreurs courantes (calcul approximatif, confusion de formules, pièges conceptuels), rendent la tâche plus difficile et discriminante (Haladyna, Downing & Rodriguez, 2002). Le choix devient alors un exercice d'analyse et non de reconnaissance.

Enfin, comme en lecture, des facteurs contextuels et logistiques peuvent également influencer la difficulté perçue d'un item mathématique : la présentation de la tâche (mise en page, clarté des illustrations), la gestion du temps, ou encore les conditions de passation (confort, bruit, stress, fatigue) peuvent altérer la performance des élèves indépendamment de leurs compétences réelles (Mislevy, Steinberg & Almond, 2003). Ces éléments doivent être pris en compte pour garantir l'équité, la validité et la comparabilité des résultats de l'évaluation.

En somme, la complexité d'une question en mathématiques ne se réduit pas à son contenu, mais découle de l'interaction entre les compétences mathématiques mobilisées, le niveau cognitif sollicité, et le contexte de présentation de la tâche. Une analyse fine de ces facteurs permet de construire des épreuves équilibrées et de mieux interpréter les performances des élèves.

1

2

3

4

2.4. Présentation des niveaux de compétence en mathématiques dans les évaluations PASEC

Pour faciliter la lecture et l'interprétation des résultats statistiques, les performances des élèves au test de mathématiques de fin de collège sont présentées sur une échelle de compétences segmentées en plusieurs niveaux.

À chaque niveau correspond un ensemble de connaissance et compétences maîtrisées, avec une certaine probabilité, par les élèves qui relèvent de ce niveau en fonction des réponses des élèves au test. Chacune des compétences et des connaissances requises à chaque niveau est présentée à la page suivante dans l'échelle de compétence PASEC2024. Ces informations sont essentielles pour apprécier le degré de maîtrise des élèves et comprendre les principales difficultés rencontrées par les élèves.

TABLEAU 7 : Échelle de compétences PASEC2024 en mathématiques – Fin de collège

Niveaux	Caractéristique du niveau	Description des compétences
Niveau 5	Résolution de problèmes complexes, modélisation, interprétation élaborée, transfert et coordination de plusieurs connaissances.	L'élève est capable de résoudre des problèmes mathématiques complexes. Cela implique de mener une démarche à plusieurs étapes, de modéliser la situation, de combiner et de croiser plusieurs relations ou grandeurs, ainsi que de mener une interprétation élaborée des résultats. En <i>Arithmétique et algèbre</i>, il résout des situations complexes impliquant de la modélisation, l'interprétation et la combinaison de plusieurs relations mathématiques. En <i>Géométrie et mesure</i>, il peut résoudre des problèmes complexes impliquant plusieurs étapes de traitement, des figures composées ou des relations métriques moins directes. En <i>Statistiques</i>, il résout un problème complexe nécessitant une interprétation élaborée ainsi qu'une sélection pertinente et critique des informations disponibles.
Niveau 4	Coordination de plusieurs procédures, choix de stratégie, interprétation de relations et organisation de la démarche.	L'élève est capable de coordonner plusieurs connaissances, propriétés ou procédures. Il sait sélectionner la méthode, la stratégie ou la représentation appropriée pour traiter, organiser et résoudre une situation mathématique simple. En <i>Arithmétique et algèbre</i>, il choisit une méthode appropriée pour traiter des équations, des relations proportionnelles, des expressions algébriques ou des fonctions simples. En <i>Géométrie</i>, il mobilise plusieurs propriétés ou représentations afin de résoudre une situation spatiale plus complexe (sélectionner une relation pertinente et organiser une démarche de résolution). En <i>Mesure</i>, il combine différentes grandeurs, figures ou formules et sélectionne une stratégie adaptée pour résoudre un problème simple. En <i>Statistique</i>, il traite ou interprète des données en comparant, calculant ou organisant des informations issues de tableaux ou de graphiques.
Niveau 3	Mobilisation et mise en relation de concepts mathématiques	L'élève est capable de mobiliser plusieurs connaissances ou procédures mathématiques dans des situations moins directes nécessitant une analyse des informations, une mise en relation de concepts ou une modélisation simple. Il sélectionne la démarche la plus appropriée parmi plusieurs procédures possibles et adapte ses connaissances à des contextes mathématiques variés. En <i>Arithmétique</i> et en <i>Algèbre</i>, il interprète et construit des expressions algébriques, manipule des expressions plus complexes, utilise des fonctions linéaires ou affines et mobilise la proportionnalité dans des situations nécessitant plusieurs étapes de raisonnement. En <i>Géométrie</i>, il applique des théorèmes, réalise ou interprète des transformations, utilise des propriétés géométriques pour déterminer des données manquantes et met en œuvre des raisonnements déductifs dans des configurations planes ou spatiales. En <i>Mesure</i>, il résout des problèmes impliquant des figures composées ou des solides variés, choisit et combine des formules adaptées et interprète les résultats obtenus dans leur contexte. Les tâches proposées exigent une compréhension plus approfondie des relations mathématiques, une certaine autonomie dans le choix des procédures et une capacité à relier différentes notions au sein d'une même situation-problème.

Niveaux	Caractéristique du niveau	Description des compétences
Seuil « suffisant » de compétences		
Niveau 2	Maîtrise fondamentale des connaissances et procédures	L'élève est capable de mobiliser des connaissances, procédures et formules mathématiques fondamentales dans des situations directes ou familières. Il réalise des calculs numériques et algébriques simples, applique des formules usuelles de mesure, interprète des représentations graphiques courantes et résout des problèmes dont la démarche est explicite ou immédiatement identifiable. En <i>Arithmétique</i> et en <i>Algèbre</i>, il utilise des expressions algébriques simples pour représenter une situation, effectue des calculs de proportionnalité, développe ou réduit des expressions élémentaires et applique des relations numériques usuelles. En <i>Géométrie</i> et en <i>Mesure</i>, il calcule des périmètres, des aires et des volumes de figures ou de solides usuels, applique des propriétés géométriques simples, détermine des coordonnées ou des distances dans le plan et utilise des relations métriques connues lorsque les données nécessaires sont directement accessibles. En <i>Statistiques</i>, il extrait et interprète des informations à partir de tableaux, diagrammes ou graphiques courants. Les tâches proposées nécessitent essentiellement l'application correcte d'une procédure connue ou d'une formule appropriée dans un contexte peu complexe.
Niveau 1	Reconnaissance directe, calculs simples, lecture immédiate de données, utilisation visible de propriétés géométriques ou de relations numériques.	L'élève est capable de reconnaître, lire et utiliser des notions, propriétés ou informations simples et directement accessibles dans des tâches directes. Il peut effectuer un calcul élémentaire, identifier une relation évidente ou appliquer une procédure très courte dans une situation familière et visuelle. En <i>Arithmétique et algèbre</i>, il reconnaît et utilise des notions numériques ou algébriques simples dans des tâches directes (calcul direct, relation simple, procédure courte en situation familière). En <i>Géométrie</i>, il reconnaît et utilise des propriétés géométriques simples liées aux figures, aux angles, aux solides ou à des relations spatiales directement visibles. En <i>Mesure</i>, il peut utiliser des unités ou appliquer directement une formule simple lorsque les données sont explicites. En <i>Statistiques</i>, il lit et extrait directement des informations explicites dans des tableaux ou des représentations graphiques simples.

Cette échelle est présentée dans le rapport international et les rapports nationaux PASEC2024 et permet d'identifier et de comparer le pourcentage moyen d'élèves qui atteignent un certain niveau ou le seuil de compétences « suffisant » pour tous les pays confondus ; par pays ; par zone éducative dans chaque pays ; par variable d'intérêt traitée par le PASEC. Le lecteur est invité à se référer à ces rapports pour obtenir ces différents résultats.

D'autres informations méthodologiques plus poussées concernant la construction des échelles de compétences sont disponibles dans le rapport technique PASEC2024.

Chapitre 3

Processus de construction des instruments et de collecte des données

Le développement des instruments d'évaluation et leur administration sur le terrain pour la collecte des données constituent une étape importante dans la réalisation d'une évaluation à grande échelle comme celle du PASEC2024. La conception des instruments d'évaluation du PASEC repose sur les principes contemporains de la validité des mesures éducatives. Selon Messick (1995), la validité ne se limite pas à la précision des scores obtenus, mais concerne également la pertinence des interprétations qui en sont faites, la représentativité des construits évalués, les processus cognitifs mobilisés par les élèves ainsi que les conséquences de l'utilisation des résultats. Dans cette perspective, le développement des instruments doit garantir que les items mesurent effectivement les compétences visées et que les différences observées entre les élèves reflètent leurs acquis plutôt que des facteurs extérieurs au construit évalué.

Cette exigence revêt une importance particulière dans les évaluations internationales, où les instruments sont administrés dans des contextes linguistiques, culturels et éducatifs très diversifiés. Le développement des épreuves du PASEC intègre ainsi des procédures rigoureuses de traduction, d'adaptation culturelle et de validation afin de réduire les biais susceptibles d'affecter la comparabilité des résultats entre les pays. L'objectif est de garantir que les différences de performance observées traduisent effectivement des différences de compétences et non des variations liées aux formulations des items ou aux contextes culturels dans lesquels ils sont administrés (Hambleton et al., 2005 ; International Test Commission, 2018).

La qualité des données recueillies repose largement sur la rigueur méthodologique utilisée pour concevoir, valider et adapter les tests aux contextes des pays participants. Ce processus englobe la conception des items, leur validation par des experts en éducation et en psychométrie, ainsi que la mise à l'essai pour garantir la pertinence, la validité, la fiabilité des instruments d'évaluation, et enfin la collecte finale.

Les progrès de la psychométrie moderne, et plus particulièrement le développement de la théorie de la réponse aux items (*Item Response Theory* – IRT), ont profondément renouvelé la conception des évaluations éducatives à grande échelle. Contrairement à la théorie classique des tests, l'IRT permet d'estimer simultanément les caractéristiques psychométriques des items (notamment leur difficulté, leur pouvoir discriminant et, selon le modèle retenu, le paramètre de pseudo-chance) ainsi que le niveau de compétence des élèves sur une échelle commune (Embretson & Reise, 2000 ; de Ayala, 2009). Cette approche offre la possibilité de construire des épreuves couvrant un large continuum de compétences tout en garantissant une mesure précise pour des élèves présentant des niveaux de performance variés.

L'un des principaux atouts de l'IRT réside dans sa capacité à assurer la comparabilité des résultats entre différentes versions d'une même épreuve grâce au recours à des items d'ancrage (*anchor items*) et aux procédures d'équating. Ces méthodes permettent de placer les performances des élèves sur une même échelle de compétence, même lorsque tous les participants ne répondent pas exactement aux mêmes items. Elles constituent ainsi un fondement méthodologique essentiel des évaluations comparatives internationales, telles que le PASEC, qui doivent

garantir la comparabilité des résultats entre les pays participants ainsi qu'entre les différents cycles d'évaluation (De Ayala, 2009 ; Embretson & Reise, 2000 ; OCDE, 2023 ; Rutkowski et al., 2014).

Le développement des instruments du PASEC2024 s'appuie sur ces principes psychométriques ainsi que sur les standards internationaux relatifs à la conception des évaluations éducatives. Le processus comprend plusieurs étapes complémentaires, allant de l'élaboration du cadre de référence et de la rédaction des items à leur validation conceptuelle, linguistique et psychométrique, avant leur intégration dans les instruments définitifs de collecte de données.

La présente partie décrit les principales étapes de ce processus. Elle présente successivement les modalités de conception et de validation des items, les procédures d'adaptation linguistique et culturelle, les analyses psychométriques réalisées lors des essais pilotes, les principes de construction des cahiers de test selon l'approche des cahiers tournants (*rotated booklet design*), ainsi que les modalités d'administration des épreuves et de collecte des données. L'ensemble de ces procédures vise à garantir la validité, la fiabilité, l'équité et la comparabilité des mesures produites par le PASEC2024.

3.1. Conception des tests

La conception des tests dans le cadre du PASEC2024 repose sur une approche rigoureuse et participative, impliquant les pays participants à toutes les étapes de développement des items. Cette approche vise non seulement à garantir la pertinence culturelle et contextuelle des items, mais aussi à renforcer les capacités techniques des pays en matière de développement d'évaluations à grande échelle.

3.1.1. Approche participative dans la conception des tests

Contrairement aux approches centralisées souvent utilisées dans les évaluations internationales, le PASEC adopte une démarche collaborative dans la production des items. Les pays participants jouent un rôle actif dans la proposition et la révision des questions, sous la supervision et avec l'accompagnement technique des experts du PASEC.

La littérature en évaluation éducative souligne que la qualité d'une épreuve internationale repose non seulement sur ses propriétés psychométriques, mais également sur la représentativité des construits évalués et sur l'adéquation des items aux contextes dans lesquels ils sont administrés. L'implication d'experts nationaux favorise ainsi le développement d'items culturellement pertinents, tout en réduisant les risques de biais liés aux différences linguistiques ou contextuelles susceptibles d'affecter l'interprétation des résultats (Hambleton et al., 2005 ; International Test Commission [ITC], 2017 ; Messick, 1995). En impliquant les experts nationaux dans le processus de création des items, le PASEC veille à ce que les tests soient adaptés aux contextes éducatifs spécifiques de chaque pays, tout en respectant les standards internationaux en matière d'évaluation.

Chaque pays est ainsi invité à proposer des items en lecture et en mathématiques en s'appuyant sur le cadre de référence commun du PASEC et sur les spécifications techniques définies pour chaque domaine d'évaluation. Les propositions sont ensuite soumises à plusieurs cycles de révision impliquant des spécialistes des disciplines concernées, des psychométriciens et des experts internationaux de l'évaluation. Ces revues successives permettent de vérifier la conformité des items aux construits évalués, leur qualité rédactionnelle, leur adéquation curriculaire ainsi que leur potentiel psychométrique, avant leur intégration dans les phases d'essais pilotes.

Une attention particulière est accordée à l'équivalence linguistique et culturelle des instruments. Les procédures de traduction, d'adaptation et de révision sont mises en œuvre conformément aux recommandations internationales afin de préserver le sens des items et de garantir que les différences observées entre les pays reflètent principalement des différences de compétence plutôt que des variations de formulation ou de contexte culturel (Hambleton et al., 2005 ; ITC, 2017). Cette démarche contribue à renforcer la validité, l'équité et la comparabilité internationale des mesures produites par le PASEC.

3.1.2. Accompagnement technique et renforcement des capacités

Le PASEC accompagne également les pays participants dans le développement technique des items à travers des ateliers de formation et des sessions de renforcement des capacités. Cet accompagnement technique vise à doter les équipes nationales des compétences nécessaires pour concevoir des items de qualité, fondés sur les modèles psychométriques modernes, comme la théorie de la réponse aux items (IRT). Les formations incluent des modules sur la rédaction d'items, la définition des niveaux de difficulté, et la construction de tests équilibrés qui couvrent les différentes compétences ciblées (Embretson & Reise, 2000). Cet effort de renforcement des capacités est crucial pour assurer une autonomie croissante des pays dans le développement de leurs propres évaluations nationales, tout en respectant les standards internationaux.

3.1.3. Processus de révision et mise à l'essai des items

Une fois les items produits et révisés par les équipes nationales et internationales, ils sont mis à l'essai dans plusieurs pays. Cette mise à l'essai est essentielle pour évaluer la pertinence des questions et leur fonctionnement dans des contextes éducatifs variés. Les résultats de la mise à l'essai permettent d'identifier les items problématiques – par exemple, des questions trop difficiles, mal comprises ou mal formulées – et d'apporter les ajustements nécessaires avant la mise en œuvre de l'évaluation à grande échelle.

Le processus de conception des tests du PASEC est donc itératif, combinant expertise locale et internationale. Ce modèle de développement est conforme aux recommandations de la littérature en évaluation éducative, qui souligne l'importance d'associer les

acteurs nationaux à la conception des instruments afin de renforcer leur pertinence culturelle, leur validité de contenu et leur acceptabilité auprès des systèmes éducatifs concernés (Greaney & Kellaghan, 2008 ; ITC, 2017 ; Messick, 1995). En favorisant la contextualisation des instruments sans compromettre les exigences de standardisation, cette approche contribue à renforcer la validité, l'équité et la comparabilité des résultats produits par le PASEC.

3.2. Validation des instruments

La validation des instruments d'évaluation constitue une étape essentielle du processus de développement des tests du PASEC2024. Elle vise à recueillir un ensemble de preuves permettant de démontrer que les scores obtenus constituent une mesure valide, fiable et équitable des compétences que l'évaluation cherche à apprécier. Dans la perspective contemporaine de la validité, il ne s'agit pas uniquement de vérifier la qualité des items pris individuellement, mais de s'assurer que les interprétations et les utilisations des résultats reposent sur des fondements scientifiques solides (*American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME]*, 2014 ; Messick, 1995).

Dans cette optique, le PASEC utilise une approche rigoureuse et en plusieurs phases, intégrant à la fois des méthodes qualitatives et quantitatives pour la validation des items.

Le processus de validation mis en œuvre par le PASEC2024 repose sur une démarche progressive et multidimensionnelle combinant des approches qualitatives et quantitatives. Les différentes étapes visent notamment à vérifier la pertinence des construits évalués, la qualité rédactionnelle des items, leur adéquation aux cadres de référence, leur adaptation linguistique et culturelle, ainsi que leurs propriétés psychométriques. Les informations recueillies au cours de ces différentes phases permettent d'améliorer progressivement les instruments et de s'assurer qu'ils mesurent effectivement les compétences visées dans l'ensemble des pays participants. Cette démarche de validation contribue également à identifier et à corriger les sources potentielles d'erreur de mesure ou de biais susceptibles d'affecter l'interprétation des résultats. Elle constitue ainsi un élément fondamental pour garantir la validité, la fiabilité, l'équité et la comparabilité internationale des évaluations réalisées dans le cadre du PASEC2024.

3.2.1. Validation qualitative : le laboratoire cognitif

L'une des spécificités du processus de validation du PASEC réside dans l'utilisation d'une validation qualitative initiale des items, réalisée à travers un laboratoire cognitif. Ce processus, inspiré des méthodologies issues de la psychologie cognitive et de la psychométrie, permet d'expérimenter les items auprès d'élèves avant leur mise en œuvre à grande échelle.

Le laboratoire cognitif consiste à soumettre les élèves cibles à des tests expérimentaux contenant des items en cours de développement, tout en observant leurs comportements et leurs réactions

1

2

3

4

face aux questions. Selon les recherches en psychométrie, les tests cognitifs permettent d'examiner comment les élèves interprètent et comprennent les items, offrant ainsi des informations cruciales pour l'ajustement des items avant leur phase finale (Willis, 2005). Cette expérimentation est suivie d'entretiens approfondis avec les élèves, au cours desquels ils partagent leurs ressentis et leurs difficultés face aux items.

Ces entretiens permettent aux concepteurs des tests de recueillir des données qualitatives sur la clarté des items, leur pertinence culturelle, ainsi que la complexité perçue des questions. Le laboratoire cognitif s'appuie sur des théories de la cognition qui postulent que la manière dont les élèves perçoivent et traitent les informations peut varier considérablement en fonction de leur expérience éducative et de leur contexte socioculturel (Ericsson & Simon, 1993). En observant ces interactions de près, le PASEC peut ajuster les items pour les rendre plus accessibles, éliminer les ambiguïtés, et s'assurer qu'ils sont équitables pour tous les élèves, indépendamment de leur milieu ou de leur langue.

3.2.2. Ajustements et révisions des items

À l'issue du laboratoire cognitif, les informations recueillies sont analysées afin d'identifier les difficultés rencontrées par les élèves lors de la compréhension et de la résolution des items. Les observations réalisées permettent de mettre en évidence d'éventuelles ambiguïtés dans les consignes, des formulations susceptibles d'être interprétées différemment selon les contextes linguistiques ou culturels, ainsi que des difficultés liées aux supports, aux illustrations ou aux situations proposées. Les résultats de cette phase servent à orienter les ajustements nécessaires avant la mise à l'essai des instruments.

Les révisions peuvent notamment porter sur la formulation des questions, le choix du vocabulaire, la clarté des consignes, la pertinence des contextes mobilisés, les illustrations, le niveau de difficulté attendu ou encore la structure des réponses proposées. L'objectif est de s'assurer que chaque item mesure effectivement le construit visé et que les réponses des élèves reflètent leurs compétences plutôt que des difficultés de compréhension ou des facteurs extérieurs au construit évalué (AERA, APA, & NCME, 2014 ; Messick, 1995).

Les ajustements réalisés à l'issue du laboratoire cognitif contribuent à réduire les risques de biais liés à la traduction, aux références culturelles ou aux contextes de passation susceptibles d'affecter la comparabilité des résultats entre les pays participants. Les recommandations internationales soulignent en effet que l'adaptation des instruments doit préserver le construit évalué tout en garantissant que les items soient compris de manière équivalente par les différentes populations concernées (Hambleton et al., 2005 ; ITC, 2017).

L'intégration des résultats qualitatifs dans le processus de révision contribue ainsi à renforcer la validité de contenu des instruments, leur équité et leur comparabilité internationale. Cette démarche permet de limiter les sources potentielles de biais et d'améliorer progressivement la qualité des items avant leur évaluation psychométrique lors des essais pilotes.

3.2.3. Validation quantitative : pré-test et analyses psychométriques

Après les ajustements qualitatifs issus du Laboratoire cognitif, les items sont ensuite soumis à une validation quantitative, notamment à travers un pré-test à grande échelle. Le pré-test permet de collecter des données quantitatives sur la performance des élèves face aux items et d'analyser statistiquement leur pertinence, leur difficulté, et leur discrimination. Ces analyses sont réalisées en utilisant le modèle IRT, qui permet d'évaluer la qualité des items de manière fine et objective (Embretson & Reise, 2000).

L'IRT est particulièrement utile dans les évaluations internationales comme le PASEC, car elle permet de s'assurer que les items fonctionnent de manière équitable pour des élèves provenant de contextes très différents. Les analyses psychométriques permettent de déterminer si les items mesurent les compétences de manière fiable à travers différents niveaux de performance, et d'identifier les items qui ne fonctionnent pas de manière optimale, c'est-à-dire ceux qui présentent des biais ou des anomalies statistiques (Wu, 2010).

Ainsi, la validation des instruments d'évaluation dans le cadre du PASEC2024 repose sur un processus rigoureux et itératif, combinant des approches qualitatives et quantitatives. Grâce à l'utilisation du Laboratoire cognitif, qui permet une expérimentation en amont des items et des ajustements en fonction des retours des élèves, le PASEC s'assure que ses tests sont non seulement fiables et valides, mais aussi culturellement adaptés et équitables. Ce processus de validation est essentiel pour garantir la qualité des données collectées et pour permettre une interprétation correcte des résultats à l'échelle internationale.

3.3. Structure et montage final des épreuves PASEC2024 : méthode des cahiers tournants

Les évaluations internationales à grande échelle sont confrontées à une contrainte majeure : couvrir un large éventail de domaines et de compétences sans imposer aux élèves un temps de passation excessif. Pour répondre à cette exigence, le PASEC2024 recourt à la méthode des cahiers tournants (*rotated booklet design*), également connue dans la littérature sous les appellations *Balanced Incomplete Block Design* (BIBD) ou *Multiple Matrix Sampling* (MMS) (Mislevy, 1991 ; OECD, 2023).

La méthode des « cahiers tournants », utilisée dans de nombreux programmes d'évaluation des acquis des élèves à l'échelle internationale, consiste à diviser le test en plusieurs blocs d'items. Chaque sous-échantillon d'élèves est alors affecté à un sous-ensemble de ces blocs, contenus dans un livret spécifique. L'avantage de cette approche réside dans l'utilisation des modèles de la réponse aux items qui permet de rapporter les performances des élèves sur différents livrets de contenu tout en utilisant une échelle de mesure commune. Cela garantit la comparabilité des scores obtenus, même si les élèves n'ont pas répondu aux mêmes items. Des blocs d'items d'ancrage sont utilisés pour lier tous les livrets entre eux, afin de garantir que les items soient positionnés sur une même échelle de difficulté et de compétence et de prédire

les performances des élèves pour les items auxquels ils n'ont pas été exposés (Hambleton et al., 1991 ; Grapin et Mounier, 2019).

Dans cette méthode, chaque bloc d'items est conçu pour couvrir les mêmes domaines d'évaluation et présenter un niveau de difficulté comparable. Les items sont organisés de manière à suivre une progression de difficulté, ce qui permet d'évaluer les élèves sur une large gamme de compétences tout en minimisant les biais liés à l'apprentissage préalable et à la fatigabilité des élèves. Ce procédé de distribution des items entre les livrets est un moyen d'assurer une distribution équitable des questions, et d'optimiser l'équité de l'évaluation, en neutralisant les effets potentiels liés à l'ordre de passation (Birnbbaum, 1968).

Grâce à la flexibilité de l'IRT, cette méthode permet d'augmenter le nombre d'items tout en maintenant une durée de passation raisonnable pour chaque élève, ce qui conduit à une meilleure précision des mesures sans alourdir le temps de test. Concrètement, chaque élève ne répond qu'à une partie du test, mais l'augmentation du nombre d'items dans le test global permet de renforcer la validité des résultats. Le test de lecture, par exemple, est constitué de 98 items répartis sur quatre blocs dans quatre livrets différents. De manière similaire, le test de mathématiques comprend 84 items répartis de façon identique. Chaque livret contient une combinaison différente des blocs, et chaque élève répond donc à environ 46 items en lecture et 40 en mathématiques, selon le livret qui lui est attribué.

Les blocs sont organisés de façon qu'ils apparaissent de manière égale dans chaque livret, ce qui permet de s'assurer que les élèves sont évalués sur les mêmes critères fondamentaux tout en ayant des sous-ensembles d'items différents. Ce système optimise la gestion du temps et la fiabilité des résultats obtenus, car chaque élève répond à des ensembles comparables d'items tout en évitant la répétition et les effets d'ennui ou de fatigue.

3.4. Administration des tests PASEC2024

L'administration des tests est également organisée de manière à minimiser les biais liés à l'ordre des tests. À cet effet, un ordre de passation aléatoire est attribué à chaque école, de sorte qu'une moitié des écoles commence par le test de lecture et l'autre moitié par le test de mathématiques. Cette organisation permet de neutraliser les effets d'ordre, qui pourraient influencer les performances des élèves et fausser les résultats, notamment en termes de fatigabilité et de motivation (Green, 2014).

Dans chaque école, un échantillon de 30 élèves est sélectionné aléatoirement pour participer à l'enquête. Les tests sont administrés sur deux jours consécutifs en matinée, et chaque test dure environ deux heures, avec une pause de 10 à 20 minutes après chaque session d'une heure. Chaque session de test correspond à un bloc d'items distinct. Afin de garantir que les élèves ne puissent pas anticiper les questions du bloc suivant, chaque bloc est scellé par des autocollants avant d'être distribué. En outre, l'administrateur de test joue un rôle de relance auprès des élèves qui pourraient être bloqués sur une question ou un texte, sans toutefois intervenir pour expliquer le contenu ou guider les élèves dans leurs réponses.

Le test est administré de manière autonome, sans interventions spécifiques de l'administrateur à l'exception de la présentation de deux exemples au début de chaque session pour s'assurer que tous les élèves comprennent le format des questions, à savoir qu'il n'y a qu'une seule réponse correcte pour chaque item. Ce modèle, inspiré de pratiques dans d'autres évaluations internationales (OECD, 2025), garantit la rigueur et la validité des tests tout en respectant le cadre temporel imparti à chaque élève.

1

2

3

4

Chapitre 4

Pistes d'amélioration et perspectives

La mise en œuvre du PASEC2024 constitue une étape importante dans le processus d'amélioration continue des évaluations internationales des acquis scolaires. Au-delà de la production d'indicateurs sur les performances des élèves, chaque cycle d'évaluation fournit des informations précieuses permettant d'apprécier la pertinence du cadre de référence, la qualité des instruments de mesure, les propriétés psychométriques des items ainsi que l'efficacité des procédures de collecte et de traitement des données. Cette démarche d'évaluation réflexive contribue à renforcer progressivement la qualité scientifique et méthodologique du programme.

Les cadres internationaux d'évaluation recommandent que chaque cycle fasse l'objet d'une analyse approfondie afin d'identifier les forces et les limites des dispositifs mis en œuvre, de documenter les difficultés rencontrées lors de l'administration des épreuves et d'examiner le fonctionnement des instruments dans les différents contextes nationaux. Ces analyses permettent d'orienter les révisions des cadres d'évaluation, d'améliorer les instruments et d'adapter les méthodologies aux évolutions des systèmes éducatifs et des connaissances scientifiques (AERA, APA, & NCME, 2014 ; OCDE, 2023). Dans cette perspective, les enseignements tirés du PASEC2024 contribueront à faire évoluer les futurs cycles d'évaluation, tant sur le plan conceptuel que méthodologique. Ils permettront notamment d'enrichir les cadres de référence, d'améliorer les procédures de développement et de validation des items, de renforcer la qualité psychométrique des instruments et d'intégrer les avancées récentes de la recherche en évaluation éducative. Cette dynamique d'amélioration continue constitue l'un des fondements de la crédibilité scientifique, de la pertinence et de la comparabilité internationale des évaluations conduites par le PASEC.

Les systèmes éducatifs évoluent continuellement sous l'effet des transformations sociales, économiques, technologiques et environnementales. Parallèlement, les attentes à l'égard des élèves se sont progressivement élargies pour inclure, au-delà de la maîtrise des savoirs disciplinaires, le développement de compétences permettant de s'adapter à des contextes complexes, de résoudre des problèmes nouveaux, de collaborer efficacement et de faire preuve d'esprit critique. Les cadres internationaux récents soulignent ainsi l'importance croissante des compétences transversales, de la pensée critique, de la créativité, de la résolution de problèmes complexes ainsi que des compétences numériques et socio-émotionnelles dans la préparation des jeunes aux défis du XXI^e siècle (OCDE, 2019, 2023 ; UNESCO, 2024).

Dans ce contexte, les dispositifs d'évaluation à grande échelle sont appelés à évoluer afin de demeurer en adéquation avec les transformations des curricula et les nouvelles attentes des systèmes éducatifs. Sans remettre en cause leur mission première de mesure des apprentissages fondamentaux, ils sont progressivement conduits à intégrer de nouvelles dimensions susceptibles d'enrichir la compréhension des acquis des élèves, notamment les compétences de raisonnement, les compétences numériques, les compétences socio-émotionnelles, les compétences collaboratives ou encore la capacité à mobiliser les connaissances dans des situations authentiques et interdisciplinaires (OCDE, 2023 ; UNESCO, 2024).

Pour le PASEC, ces évolutions représentent autant d'opportunités que de défis. Toute évolution du cadre d'évaluation doit préserver la comparabilité des résultats entre les cycles tout en permettant une adaptation progressive aux nouvelles orientations des politiques éducatives et aux priorités des pays membres de la CONFEMEN. Les innovations introduites devront ainsi concilier continuité méthodologique, validité des mesures et prise en compte des compétences émergentes jugées essentielles à la réussite scolaire, à l'insertion professionnelle et à l'exercice d'une citoyenneté active.

La présente section examine les principales pistes d'amélioration susceptibles d'être envisagées à la lumière des enseignements tirés des cycles précédents du PASEC ainsi que des évolutions récentes de la recherche en évaluation éducative. Elle s'intéresse notamment aux perspectives d'évolution du cadre de référence, aux innovations méthodologiques, à l'intégration progressive de nouveaux domaines de compétences et au renforcement des procédures de développement et de validation des instruments. Ces réflexions visent à consolider la capacité du PASEC à produire des informations scientifiquement robustes, comparables et utiles au pilotage des systèmes éducatifs en Afrique subsaharienne.

4.1. Réflexions sur l'évolution des cadres d'évaluation

L'évolution des cadres d'évaluation constitue une condition essentielle pour maintenir la pertinence des évaluations éducatives dans un contexte de transformation rapide des systèmes d'enseignement. Les réformes curriculaires, les avancées de la recherche sur les apprentissages ainsi que les mutations sociales, économiques, technologiques et environnementales conduisent progressivement à redéfinir les compétences que les élèves sont appelés à développer tout au long de leur scolarité. En conséquence, les cadres d'évaluation doivent évoluer afin de continuer à mesurer des compétences en adéquation avec les finalités actuelles de l'éducation tout en préservant la comparabilité des résultats dans le temps (AERA, APA, & NCME, 2014 ; OCDE, 2023).

Au cours des deux dernières décennies, la littérature scientifique et les organisations internationales ont souligné la nécessité de dépasser une conception de l'évaluation centrée exclusivement sur les connaissances disciplinaires pour intégrer des compétences plus complexes mobilisées dans des situations variées. Ces compétences, souvent désignées comme des compétences transversales, compétences globales ou compétences du XXI^e siècle, regroupent notamment la pensée critique, la créativité, la communication, la collaboration, la résolution de problèmes complexes, les compétences numériques ainsi que les compétences sociales et émotionnelles (OCDE, 2019, 2023 ; UNESCO, 2024).

Ces évolutions ne remettent pas en cause l'importance des apprentissages fondamentaux en lecture, en mathématiques et en sciences, qui demeurent le socle de toute réussite scolaire. Elles conduisent toutefois à considérer ces apprentissages dans une perspective plus large, où les élèves sont appelés non seulement à maîtriser des connaissances, mais également à les mobiliser de manière pertinente, critique et créative dans des situations nouvelles ou complexes. Les cadres d'évaluation sont

ainsi progressivement amenés à mieux prendre en compte les compétences de raisonnement, de transfert des apprentissages et de résolution de problèmes authentiques, tout en intégrant les évolutions des curricula nationaux et les priorités éducatives internationales (OCDE, 2023 ; UNESCO, 2024).

Dans cette perspective, le PASEC est appelé à faire évoluer progressivement son cadre de référence afin de continuer à fournir aux pays membres de la CONFEMEN des informations pertinentes sur les acquis des élèves et sur leur préparation aux exigences des sociétés contemporaines. Cette évolution devra toutefois préserver les principes fondamentaux qui caractérisent les évaluations internationales, notamment la validité des mesures, l'équité des instruments et la comparabilité des résultats entre les pays et entre les différents cycles d'évaluation.

4.1.1. L'évolution vers des évaluations plus complètes et diversifiées

Au cours des dernières décennies, les cadres d'évaluation des apprentissages ont connu une évolution importante sous l'influence des avancées de la psychologie cognitive, des sciences de l'apprentissage et de la didactique des disciplines. Les premières évaluations standardisées étaient principalement centrées sur la mesure des connaissances déclaratives et de la maîtrise des procédures, en accord avec les conceptions dominantes de l'apprentissage de l'époque. La taxonomie des objectifs éducatifs proposée par Bloom et al. (1956) a largement contribué à structurer ces premières approches en distinguant différents niveaux de processus cognitifs, allant de la connaissance à l'évaluation.

Les travaux ultérieurs ont toutefois montré que la réussite scolaire ne repose pas uniquement sur la mémorisation des connaissances ou l'application de procédures routinières. Les recherches en psychologie cognitive soulignent l'importance de processus cognitifs plus élaborés, tels que l'analyse, l'interprétation, le raisonnement, la résolution de problèmes, la métacognition et la créativité, qui permettent aux élèves de mobiliser leurs connaissances dans des situations variées et complexes. La révision de la taxonomie de Bloom réalisée par Anderson et Krathwohl (2001) traduit cette évolution en mettant davantage l'accent sur les processus cognitifs mobilisés par les apprenants et sur la construction active des connaissances.

Ces avancées ont progressivement influencé les cadres conceptuels des grandes évaluations internationales. Les programmes tels que PISA, TIMSS et PIRLS ne se limitent plus à mesurer la maîtrise de connaissances disciplinaires ; ils évaluent également la capacité des élèves à mobiliser ces connaissances dans des situations authentiques, à raisonner, à résoudre des problèmes et à transférer leurs apprentissages dans des contextes nouveaux (IEA, 2021 ; OCDE, 2023). Cette évolution traduit une conception plus fonctionnelle des compétences, dans laquelle les savoirs sont considérés comme des ressources à mobiliser plutôt que comme de simples connaissances à restituer.

Parallèlement, les organisations internationales mettent de plus en plus l'accent sur le développement de compétences transversales ou globales, telles que la pensée critique, la créativité, la

communication, la collaboration, les compétences numériques et les compétences socio-émotionnelles. Les recherches montrent que ces compétences complètent les apprentissages fondamentaux et jouent un rôle déterminant dans la réussite scolaire, l'insertion professionnelle et la capacité des individus à s'adapter à des environnements complexes et en constante évolution (OCDE, 2019, 2023 ; Trilling & Fadel, 2009 ; UNESCO, 2024). Les cadres d'évaluation sont ainsi appelés à évoluer afin de rendre compte de cette conception élargie des apprentissages tout en préservant la validité, l'équité et la comparabilité des mesures.

4.1.2. Intégration des compétences transversales et socio-émotionnelles

Les cadres d'évaluation modernes cherchent à intégrer des compétences transversales, telles que la pensée critique et la créativité, dans les tests académiques. Cela répond à une demande croissante de la société pour des apprenants capables de s'adapter à un monde professionnel de plus en plus complexe et imprévisible (*World Economic Forum*, 2016). Dans cette perspective, plusieurs recherches indiquent que les compétences cognitives supérieures, comme la résolution de problèmes complexes et la pensée critique, devraient être au cœur des nouvelles approches d'évaluation (Zohar & Dori, 2003).

Outre les compétences cognitives, il existe un intérêt croissant pour l'intégration des compétences socio-émotionnelles dans les cadres d'évaluation. Les études montrent que les compétences socio-émotionnelles telles que la persévérance, la gestion du stress et la capacité à travailler en équipe, sont des indicateurs clés de la réussite future des élèves, aussi bien sur le plan académique que dans leur vie personnelle et professionnelle (Duckworth & Yeager, 2015). Ces compétences, longtemps négligées dans les évaluations à grande échelle, commencent à être reconnues comme des dimensions importantes à évaluer pour fournir une image plus complète des acquis des élèves. Le PISA, par exemple, a récemment introduit des volets consacrés à la résolution collaborative de problèmes et à la gestion des émotions dans son cadre d'évaluation.

4.1.3. Évolutions liées aux technologies numériques

L'évolution des cadres d'évaluation est également influencée par l'intégration croissante des technologies numériques dans l'éducation. L'évaluation des compétences numériques et la capacité à utiliser les outils technologiques sont devenues des aspects essentiels des cadres d'évaluation contemporains (Redecker, 2013). Des environnements numériques d'évaluation permettent non seulement de mesurer les compétences en littératie numérique, mais aussi d'introduire des items interactifs qui évaluent des compétences plus dynamiques et pratiques, comme la résolution de problèmes dans des environnements numériques.

Les technologies numériques offrent également des opportunités pour rendre les évaluations plus adaptatives, c'est-à-dire capables de s'ajuster en temps réel au niveau de compétence de l'élève. Cela permet d'obtenir une mesure plus précise des

compétences individuelles et d'évaluer des dimensions cognitives qui sont difficilement mesurables avec des tests papier-crayon classiques (Shute & Rahimi, 2017).

4.2. Propositions pour l'évolution du cadre d'évaluation du PASEC

L'amélioration des processus cognitifs dans les évaluations à grande échelle, tels que ceux du PASEC, est une priorité pour mieux capturer la diversité des compétences des élèves et mieux comprendre les mécanismes d'apprentissage. De nombreuses recherches en sciences cognitives et en éducation soulignent l'importance d'évaluer des processus cognitifs plus complexes et de diversifier les types d'items pour une mesure plus nuancée des acquis des élèves (Anderson & Krathwohl, 2001 ; Pellegrino, Chudowsky & Glaser, 2001).

Dans cette perspective, les cadres internationaux d'évaluation accordent une importance croissante à l'équilibre entre différents niveaux de complexité cognitive ainsi qu'à la diversification des situations proposées aux élèves. Les évolutions récentes de PISA et de TIMSS illustrent cette orientation en privilégiant davantage des tâches mobilisant le raisonnement, la modélisation, le transfert des connaissances et la résolution de problèmes dans des contextes authentiques (IEA, 2021 ; OCDE, 2023). Cette évolution permet d'obtenir une représentation plus complète des compétences des élèves et de mieux comprendre les processus qui sous-tendent leurs apprentissages.

Pour les futurs cycles du PASEC, plusieurs pistes d'amélioration peuvent ainsi être envisagées afin d'enrichir les processus cognitifs évalués, de diversifier les formats de tâches proposés et de mieux prendre en compte la complexité croissante des compétences attendues dans les curricula contemporains. Ces évolutions devront toutefois préserver les principes fondamentaux de validité, de fiabilité, d'équité et de comparabilité qui caractérisent les évaluations internationales à grande échelle (AERA et al., 2014).

4.2.1. Renforcer l'évaluation de la résolution de problèmes complexes

Les compétences de résolution de problèmes complexes sont de plus en plus reconnues comme des leviers pour préparer les élèves aux défis du XXI^e siècle (Saavedra & Opfer, 2012). Le renforcement de l'évaluation de la résolution de problèmes complexes constitue l'une des principales orientations des évaluations éducatives contemporaines. Dans un contexte marqué par l'accélération des transformations technologiques, économiques et sociales, les élèves sont de plus en plus amenés à mobiliser leurs connaissances dans des situations inédites, à raisonner de manière critique, à prendre des décisions fondées sur des informations parfois incomplètes et à élaborer des solutions adaptées à des problèmes multidimensionnels. Les organisations internationales reconnaissent désormais ces compétences comme essentielles à la réussite scolaire, à l'apprentissage tout au long de la vie et à l'insertion professionnelle (OCDE, 2019, 2023 ; UNESCO, 2024).

Cependant, de nombreuses évaluations internationales et nationales se concentrent encore principalement sur des tâches routinières ou des problèmes simples, qui ne nécessitent pas la mobilisation de plusieurs ressources ou stratégies. Pour améliorer l'évaluation de ces compétences, il est essentiel de développer des items qui sollicitent les processus cognitifs supérieurs, tels que l'analyse, la synthèse et l'évaluation (Anderson & Krathwohl, 2001 ; Bloom et al., 1956).

Les travaux de Jonassen (2000) distinguent notamment les problèmes mal définis (*Ill-structured problems*), caractérisés par l'absence d'une procédure unique de résolution, la présence de plusieurs solutions plausibles et la nécessité de mobiliser un jugement critique. Ce type de tâches favorise le développement du raisonnement, de la créativité et de la capacité à transférer les connaissances dans des contextes nouveaux. Leur intégration progressive dans les évaluations à grande échelle permettrait d'obtenir une représentation plus complète des compétences des élèves et de mieux refléter les exigences des curricula contemporains.

Les grandes évaluations internationales évoluent progressivement dans cette direction. PISA a ainsi introduit des domaines innovants consacrés à la résolution de problèmes, à la résolution collaborative de problèmes, à la pensée créative et au raisonnement mathématique, traduisant une volonté de mesurer la capacité des élèves à mobiliser leurs connaissances dans des situations authentiques et complexes (OCDE, 2017, 2023). Ces évolutions montrent qu'il est possible d'intégrer des tâches plus exigeantes sur le plan cognitif tout en préservant la validité psychométrique et la comparabilité internationale des résultats.

Dans cette perspective, les futurs cycles du PASEC pourraient progressivement enrichir les situations proposées aux élèves en intégrant davantage de problèmes contextualisés, multidimensionnels et ouverts, mobilisant plusieurs domaines de connaissances et différents processus cognitifs. Une telle évolution permettrait d'apprécier non seulement les connaissances acquises, mais également la capacité des élèves à les mobiliser de manière pertinente, flexible et argumentée dans des situations proches de celles rencontrées dans la vie scolaire, sociale ou professionnelle.

4.2.2. Diversifier les formats d'items pour stimuler différentes formes de pensée

Les évaluations internationales à grande échelle reposent traditionnellement sur des questions à choix multiple et des questions à réponse construite courte, qui présentent l'avantage d'assurer une correction standardisée, une bonne fidélité de mesure et une comparabilité entre les populations évaluées. Toutefois, ces formats ne permettent pas toujours d'appréhender l'ensemble des processus cognitifs mobilisés par les élèves, notamment lorsqu'il s'agit d'évaluer des compétences complexes telles que le raisonnement, la résolution de problèmes, l'argumentation ou le transfert des connaissances (AERA, APA, & NCME, 2014 ; Pellegrino et al., 2001).

Le développement des évaluations numériques ouvre également de nouvelles perspectives grâce à l'utilisation d'items multimodaux, qui associent textes, graphiques, tableaux, images, vidéos, animations ou environnements interactifs. Ces formats permettent d'évaluer la capacité des élèves à interpréter, intégrer et mobiliser des informations provenant de plusieurs sources ou représentations, compétence devenue essentielle dans les environnements numériques contemporains (Mayer & Fiorella, 2022 ; OCDE, 2023). Les recherches sur l'apprentissage multimédia montrent en effet que les élèves sont de plus en plus amenés à construire du sens à partir d'informations distribuées entre différents modes de représentation, mobilisant ainsi des processus cognitifs plus complexes que la seule lecture linéaire d'un texte ((Mayer & Fiorella, 2022).

Par ailleurs, les tâches de performance (*performance tasks*) et les scénarios complexes permettent d'évaluer la mobilisation simultanée de plusieurs connaissances et compétences dans des situations proches de celles rencontrées dans la vie scolaire ou quotidienne. Elles offrent la possibilité d'observer non seulement le résultat obtenu, mais également la démarche de résolution, la qualité du raisonnement, la justification des réponses et la capacité des élèves à sélectionner des stratégies adaptées à des situations inédites (Shavelson, 2013). Ces formats apparaissent particulièrement prometteurs pour l'évaluation du raisonnement mathématique, de la compréhension approfondie en lecture et de la résolution de problèmes complexes.

Dans cette perspective, les futurs cycles du PASEC pourraient progressivement diversifier les formats d'items utilisés, en intégrant, lorsque les conditions techniques et opérationnelles le permettront, des tâches plus authentiques, contextualisées et interactives. Une telle évolution contribuerait à enrichir les informations recueillies sur les compétences des élèves tout en maintenant les exigences de validité, de fiabilité, d'équité et de comparabilité qui caractérisent les évaluations internationales à grande échelle.

4.2.3. Intégrer l'évaluation des compétences transversales et socio-émotionnelles

Les compétences socio-émotionnelles, telles que la persévérance, la régulation émotionnelle et la collaboration, sont de plus en plus reconnues comme essentielles pour la réussite scolaire et professionnelle (Duckworth & Yeager, 2015). Ces compétences sont encore peu intégrées dans les évaluations traditionnelles, qui se concentrent majoritairement sur les compétences cognitives. Pour améliorer la pertinence des processus cognitifs évalués, il serait nécessaire d'incorporer des items qui permettent de mesurer ces compétences de manière directe ou indirecte.

Les recherches de Durlak et al. (2011) montrent que les élèves dotés de solides compétences socio-émotionnelles ont de meilleurs résultats scolaires et une plus grande capacité à surmonter les défis académiques. Le cadre du PASEC pourrait donc être enrichi par des items qui évaluent la capacité des élèves à persévérer face à des tâches difficiles, à gérer des émotions dans des situations de résolution de problèmes, ou à collaborer efficacement avec leurs pairs. Le PISA a déjà commencé à inclure des volets de « résolution collaborative de problèmes », et des innovations similaires pourraient être envisagées pour le PASEC.

4.2.4. Utiliser les technologies numériques pour évaluer des processus cognitifs dynamiques

Les avancées technologiques offrent des opportunités sans précédent pour l'amélioration des évaluations cognitives. En intégrant des outils numériques, il est possible de mesurer des compétences dynamiques qui étaient auparavant difficiles à évaluer, comme la résolution de problèmes en temps réel, l'adaptation des stratégies ou la pensée algorithmique. Les recherches de Shute et Rahimi (2017) sur les environnements d'évaluation numériques montrent que ces outils permettent de capturer des informations plus fines sur le processus cognitif des élèves, telles que la manière dont ils explorent différentes solutions ou ajustent leurs stratégies en cours de résolution.

Les environnements d'évaluation adaptatifs, par exemple, ajustent automatiquement la difficulté des questions en fonction du niveau de performance de l'élève, ce qui permet de mieux évaluer les processus cognitifs à tous les niveaux de compétence (Pellegrino & Quellmalz, 2010). De plus, l'utilisation de simulations et d'items interactifs permet d'évaluer des compétences complexes dans des contextes réalistes, tout en rendant l'expérience d'évaluation plus engageante pour les élèves.

En conclusion, pour améliorer l'évaluation des processus cognitifs dans les tests à grande échelle comme le PASEC, il est nécessaire de renforcer l'évaluation de la résolution de problèmes complexes, de diversifier les formats d'items, d'intégrer les compétences socio-émotionnelles et transversales, et d'exploiter les technologies numériques. Ces améliorations permettront de mieux mesurer les compétences cognitives et non-cognitives des élèves, tout en offrant des données plus riches pour orienter les réformes éducatives.

Conclusion

Le cadre de référence pour l'évaluation PASEC2024 en fin de collège vise à mesurer de manière rigoureuse les compétences des élèves en lecture et en mathématiques. L'évaluation a pour objectifs principaux de fournir des données comparatives et informatives sur les systèmes éducatifs des pays participants en Afrique subsaharienne. En mesurant les connaissances et les compétences acquises à un stade crucial de la scolarité, le PASEC permet d'évaluer dans quelle mesure les élèves maîtrisent les compétences fondamentales, indispensables pour leur réussite future dans l'éducation et dans la vie professionnelle. Cette approche vise également à fournir des données probantes pour guider les réformes éducatives et améliorer la qualité des systèmes éducatifs.

La mesure des processus cognitifs dans les évaluations internationales, comme le PASEC, revêt une importance capitale pour l'amélioration des systèmes éducatifs en Afrique subsaharienne. En évaluant non seulement la maîtrise des compétences de base mais, aussi la manière dont les élèves mobilisent ces compétences pour résoudre des problèmes complexes et appliquer leurs connaissances dans divers contextes, le PASEC contribue à une compréhension plus fine des processus d'apprentissage. L'évaluation de ces processus cognitifs avancés est essentielle pour identifier les forces et les faiblesses des élèves, mais aussi pour comprendre les écarts de performance entre les pays et entre les différents groupes d'élèves à l'intérieur d'un même pays.

Les perspectives ouvertes par l'évaluation PASEC2024 offrent de nouvelles opportunités pour améliorer les systèmes éducatifs dans les pays participants. Une meilleure compréhension des acquis des élèves, basée sur des données fiables et comparables, permet d'orienter les réformes éducatives de manière plus ciblée et plus efficace. En fournissant des informations détaillées sur les compétences acquises et sur les processus cognitifs sous-jacents, l'évaluation offre une base solide pour repenser les programmes scolaires, améliorer les pratiques pédagogiques et promouvoir des environnements d'apprentissage plus inclusifs et équitables. De plus, l'analyse des données issues de l'évaluation

peut aider à identifier des leviers d'action spécifiques, tels que le renforcement des compétences transversales ou l'intégration des technologies numériques dans l'enseignement.

En définitive, l'évolution des processus cognitifs évalués constitue un levier majeur pour renforcer la pertinence et la qualité des évaluations internationales à grande échelle. Les avancées récentes des sciences de l'apprentissage, de la psychométrie et des technologies éducatives offrent aujourd'hui de nouvelles possibilités pour mieux appréhender la diversité des compétences mobilisées par les élèves dans des situations scolaires et authentiques. Le renforcement de l'évaluation du raisonnement et de la résolution de problèmes complexes, la diversification des formats de tâches, l'intégration progressive de compétences transversales ainsi que le recours aux environnements numériques d'évaluation représentent autant de pistes prometteuses pour enrichir les futurs cadres d'évaluation.

Toutefois, ces évolutions devront être mises en œuvre de manière progressive et rigoureuse afin de préserver les principes fondamentaux qui sous-tendent les évaluations internationales, notamment la validité des construits évalués, la fiabilité des mesures, l'équité des instruments et la comparabilité des résultats entre les pays et entre les cycles d'évaluation (AERA, APA, & NCME, 2014 ; OCDE, 2023). Elles devront également tenir compte des réalités pédagogiques, linguistiques et technologiques des pays participants.

Dans cette perspective, le PASEC dispose d'un potentiel important pour poursuivre l'évolution de son cadre d'évaluation en intégrant progressivement les innovations méthodologiques reconnues par la recherche internationale, tout en demeurant fidèle à sa mission première de produire des données robustes, comparables et utiles au pilotage des systèmes éducatifs. Cette démarche d'amélioration continue contribuera à renforcer la capacité du PASEC à accompagner les pays membres de la CONFEMEN dans leurs efforts visant à améliorer la qualité, l'équité et la pertinence des apprentissages au regard des défis éducatifs contemporains.

Bibliographie

- Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, 61(5), 364-373.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Allyn & Bacon.
- Battista, M. T. (2007). The Development of Geometric and Spatial Thinking. Dans : F. K. Lester Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 843-908). Information Age Publishing.
- Biemiller, A. (2003). *Oral comprehension sets the ceiling on reading comprehension*. *American Educator*, 27(1), 23-29.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. Academic Press.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. Dans : F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397-479). Reading, MA : Addison-Wesley.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I: Cognitive Domain*. Longmans, Green.
- Britt, M. A., & Rouet, J.-F. (2012). Learning with Multiple Documents: Component Skills and Their Acquisition. Dans : J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes* (pp. 276-314). Cambridge University Press.
- Clements, D. H. (2004). Geometric and Spatial Thinking in Early Childhood Education. Dans : D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging Young Children in Mathematics: Standards for Early Childhood Mathematics Education* (pp. 267-297). Lawrence Erlbaum Associates.
- Cobb, G. W., & Moore, D. S. (1997). *Mathematics, statistics, and teaching*. *The American Mathematical Monthly*, 104(9), 801-823. <https://doi.org/10.1080/00029890.1997.11990723>
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement Matters: Assessing Personal Qualities Other than Cognitive Ability for Educational Purposes. *Educational Researcher*, 44(4), 237-251. <https://doi.org/10.3102/0013189X15584327>
- Duke, N. K., & Cartwright, K. B. (2021). The Science of Reading Progresses: Communicating Advances beyond the Simple View of Reading. *Reading Research Quarterly*, 56, S25-S44. <https://doi.org/10.1002/rtrq.411>
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). *The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions*. *Child Development*, 82(1), 405-432.
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A PreK-12 Curriculum Framework*. American Statistical Association.
- Giasson, J. (2000). *La compréhension en lecture*. Montréal : Chenelière/McGraw-Hill.
- Grapin, G., & Mounier, E. (2019). *Compétences en lecture et compréhension de l'écrit*. Paris : Hachette Éducation.
- Greaney, V., & Kellaghan, T. (2008). *Assessing National Achievement Levels in Education*. World Bank Publications.
- Green, A. (2014). *Exploring language assessment and testing: Language in action*. London : Routledge.
- Greer, B., & Verschaffel, L. (2007). Modelling competencies in mathematics education. In W. Blum, P. Galbraith, H-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education* (pp. 159-168). Springer.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items* (3rd ed.). New York : Routledge.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309-334.
- Hambleton, R. K., & Kanjee, A. (1995). Increasing the Validity of Cross-Cultural Assessments: Use of Improved Methods for Test Adaptations. *European Journal of Psychological Assessment*, 11(3), 147-157.
- Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. L. Erlbaum Associates.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA : Sage Publications.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and Teaching with Understanding. Dans : D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 65-97). Macmillan.
- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. Dans : F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371-404). Charlotte, NC : Information Age Publishing.
- International Test Commission. (2018). ITC Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, 18(2), 101-134. <https://doi.org/10.1080/15305058.2017.1398166>
- Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology Research and Development*, 48(4), 63-85.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.) (2001). *Adding It Up: Helping Children Learn Mathematics*. National Academy Press.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60(1), 1-64.

- Lesh, R., & Doerr, H. M. (2003). *Beyond constructivism: A models and modeling perspective on mathematics problem solving, learning, and teaching*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Mayer, R. E., & Fiorella, L. (2022). *The Cambridge handbook of multimedia learning* (Third edition ed.). Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- McNamara, D. S., & Magliano, J. (2009). Chapter 9 Toward a Comprehensive Model of Comprehension. *The Psychology of Learning and Motivation*, 51, 297-384. [https://doi.org/10.1016/S0079-7421\(09\)51009-2](https://doi.org/10.1016/S0079-7421(09)51009-2)
- Messick, S. (1995). *Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning*. *American Psychologist*, 50(9), 741-749.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus on educational evaluation: On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3-62.
- Moore, D. S. (1997). *New pedagogy and new content: The case of statistics*. *International Statistical Review*, 65(2), 123-137. <https://doi.org/10.1111/j.1751-5823.1997.tb00390.x>
- Moore, D. S., Cobb, G. W., Garfield, J., & Meeker, W. Q. (1995). *Statistics education, fin de siècle*. *The American Statistician*, 49(3), 250-260.
- Mullis, I. V. S. E., Martin, M. O. E., International Association for the Evaluation of Educational, A., Boston College, T., & Center, P. I. S. (2019). *PIRLS 2021 Assessment Frameworks*. International Association for the Evaluation of Educational Achievement.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Drucker, K. T. (2017). *PIRLS 2016 assessment framework*. Chestnut Hill, MA : TIMSS & PIRLS International Study Center, Boston College.
- Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273-304. <https://doi.org/10.1598/RRQ.46.3.4>
- National Research Council. (2001). *Adding It Up: Helping Children Learn Mathematics*. National Academies Press.
- OCDE. (2023). *PISA 2022 Assessment and Analytical Framework* (1st ed.). OECD Publishing.
- OCDE. (2017). *PISA 2015 Collaborative Problem-Solving Framework*. OECD Publishing.
- OCDE. (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*. Paris : OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- OECD. (2025). *PISA 2025 Mathematics Framework*. Paris : OECD Publishing.
- Pellegrino, J. W., & Quellmalz, E. S. (2010). Perspectives on the Integration of Technology and Assessment. *Journal of Research on Technology in Education*, 43(2), 119-134.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.) (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academy Press.
- PIRLS Framework. (2017). Dans : I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2016 Assessment Framework*. Boston College, TIMSS & PIRLS International Study Center.
- Pólya, G. (1945). *How to Solve It: A New Aspect of Mathematical Method*. Princeton University Press.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ : Lawrence Erlbaum Associates.
- Redecker, C. (2013). *The Use of ICT for the Assessment of Key Competences*. European Commission, Joint Research Centre.
- Rutkowski, L., Davier, M. v., & Rutkowski, D. (2014). *Handbook of international large-scale assessment : background, technical issues, and methods of data analysis*. CRC Press. <https://doi.org/10.1201/b16061>
- Saavedra, A. R., & Opfer, V. D. (2012). *Learning 21st-Century Skills Requires 21st-Century Teaching*. *Phi Delta Kappan*, 94(2), 8-13.
- Schneuwly, B., & Dolz, J. (2009). *Les genres scolaires et l'enseignement du français : Lire, écrire, parler*. Paris : ESF Éditeur.
- Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Academic Press.
- Shavelson, R. J. (2013). Performance Assessment. Dans : J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 145-160). SAGE Publications.
- Shute, V. J., & Rahimi, S. (2017). Review of Computer-Based Assessment for Learning in Elementary and Secondary Education. *Journal of Computer Assisted Learning*, 33(1), 1-19.
- Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA : RAND Corporation.
- Snow, C. E. (2010). Academic Language and the Challenge of Reading for Learning About Science. *Science*, 328, 450. <https://doi.org/10.1126/science.1182597>
- Sriraman, B., & English, L. D. (2010). Theories of mathematics education : seeking new frontiers. *Springer*. <https://doi.org/10.1007/978-3-642-00742-2>
- Stacey, K., & Turner, R. (Eds.). (2015). *Assessing mathematical literacy: The PISA experience*. Dordrecht : Springer.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20(4), 345-354.
- Trilling, B., & Fadel, C. (2009). *21st Century Skills: Learning for Life in Our Times*. Jossey-Bass.
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York : Academic Press.
- Van Hiele, P. M. (1986). *Structure and Insight: A Theory of Mathematics Education*. Academic Press.
- World Economic Forum. (2016). *New vision for education: Fostering social and emotional learning through technology*. Geneva : World Economic Forum.
- Wu, M. (2010). Measurement, Sampling, and Equating: The Backbone of PISA and Other International Large-Scale Assessments. Dans : M. von Davier & D. Hastedt (Eds.), *Issues and Methodologies in Large-Scale Assessments* (pp. 137-157). International Association for the Evaluation of Educational Achievement (IEA).

Les évaluations internationales PASEC : de 2014 à 2024

Évolution du dispositif, spécificités des cycles et principales innovations

Depuis la réforme méthodologique engagée au début des années 2010, le **Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) conduit des évaluations internationales comparatives à grande échelle** afin de produire des données fiables et comparables sur les apprentissages, leur environnement et l'équité des systèmes éducatifs. Les cycles PASEC2014, PASEC2019 et PASEC2024 s'inscrivent dans une même architecture méthodologique, tout en intégrant progressivement de nouvelles populations, de nouveaux champs d'analyse et de nouvelles modalités de collecte. Cette évolution répond à un double objectif : préserver la comparabilité dans le temps et accroître la capacité du dispositif à éclairer le pilotage des politiques éducatives.

1. Un socle méthodologique commun

Les trois cycles reposent sur une évaluation externe et standardisée des apprentissages, conduite selon des procédures harmonisées dans les pays participants. Les populations scolaires sont étudiées à partir d'échantillons probabilistes représentatifs, permettant d'estimer les performances au niveau national sans procéder à une enquête exhaustive.

Le dispositif combine des tests de compétences en langue d'enseignement/lecture et en mathématiques avec des questionnaires contextuels. Les scores sont positionnés sur des échelles internationales et interprétés à travers des niveaux de compétences. Les informations contextuelles recueillies auprès des différents acteurs permettent de relier les performances aux caractéristiques des élèves, des enseignants, des établissements et de leur environnement.

La comparabilité repose également sur des procédures d'assurance qualité couvrant la conception et la mise à l'essai des instruments, leur adaptation linguistique, l'échantillonnage, l'administration, le contrôle de la collecte, le traitement des données et les analyses psychométriques. Les équipes nationales PASEC sont associées à ces différentes étapes, dans une logique de renforcement durable des capacités nationales en évaluation.

2. PASEC2014 : établir la référence internationale

PASEC2014 constitue le premier cycle international standardisé du Programme. Mis en œuvre dans dix pays d'Afrique subsaharienne — Bénin, Burkina Faso, Burundi, Cameroun, Congo, Côte d'Ivoire, Niger, Sénégal, Tchad et Togo — il évalue les élèves en début de scolarité primaire (2^e année) et en fin de scolarité primaire (6^e année) en langue d'enseignement/lecture et en mathématiques. Des données contextuelles sont parallèlement recueillies sur les élèves, les enseignants, les directeurs, les classes et les écoles.

En début de primaire, les épreuves sont administrées individuellement afin d'apprécier précocement les compétences fondamentales. En fin de primaire, l'évaluation repose principalement sur des épreuves papier-crayon et des cahiers tournants. Les résultats sont analysés en termes de performance, de niveaux de compétences, d'efficacité et d'équité des systèmes éducatifs.

L'innovation structurante de PASEC2014 est la création d'une métrique internationale commune. Le cycle transforme les expériences nationales antérieures du PASEC en un dispositif comparatif harmonisé et fournit une première référence commune pour suivre les systèmes éducatifs participants dans l'espace et dans le temps. L'évaluation simultanée du début et de la fin du primaire permet en outre de documenter à la fois les premiers apprentissages et les compétences indispensables à la poursuite de la scolarité.

3. PASEC2019 : consolider les tendances et intégrer les enseignants

PASEC2019 constitue le deuxième cycle international et élargit la participation à quatorze pays. Aux dix pays de 2014 s'ajoutent le Gabon, la Guinée, Madagascar et la République démocratique du Congo. L'architecture

d'évaluation des élèves en début et en fin de primaire est maintenue, ce qui permet, pour les dix pays communs aux deux cycles, d'analyser les évolutions des performances et des disparités entre 2014 et 2019.

Les questionnaires contextuels sont également révisés afin de mieux décrire l'environnement d'apprentissage : questionnaire Élève, questionnaire Enseignant/Classe et questionnaire Directeur/École. Les procédures d'adaptation linguistique, de mise à l'essai et d'assurance qualité sont renforcées afin de sécuriser la comparabilité nationale et internationale des instruments.

La principale innovation de PASEC2019 est l'introduction d'une enquête internationale sur les enseignants du primaire. Celle-ci mesure leurs connaissances disciplinaires en compréhension de l'écrit et en mathématiques ainsi que leurs connaissances didactiques, notamment leur capacité à planifier une situation d'apprentissage et à identifier ou analyser les erreurs des élèves. Elle recueille également des informations sur la formation, l'expérience professionnelle et les conditions de travail. Le diagnostic ne porte donc plus seulement sur les acquis des élèves et leur contexte : il intègre directement les connaissances et compétences professionnelles des enseignants, avec des implications pour les politiques de formation initiale et continue.

4. PASEC2024 : élargir le champ de l'évaluation et moderniser la collecte

PASEC2024 constitue le troisième cycle international et le plus large conduit par le Programme à ce jour, avec seize pays. Il conserve les composantes du primaire nécessaires au suivi des tendances, tout en élargissant le périmètre géographique (16 pays), linguistique et méthodologique de l'évaluation.

La première innovation majeure est l'extension de l'évaluation au collège / à la fin du premier cycle du secondaire. PASEC2024 permet ainsi d'appréhender plus largement le continuum de l'éducation de base, en complément des mesures réalisées en début et en fin de primaire. Cette extension renforce notamment la contribution du dispositif au suivi de la cible 4.1.1 de l'ODD 4, qui porte sur les niveaux minimaux de compétence à différents moments de la scolarité.

PASEC2024 introduit aussi une innovation linguistique importante. Le cycle marque à la fois une évolution des langues internationales mobilisées, avec l'intégration du portugais et l'élargissement de l'anglais notamment au Nigéria, et une prise en compte accrue des langues nationales utilisées comme langues d'enseignement, particulièrement dans les premières années de scolarité. Cette double évolution permet de mieux refléter la réalité des contextes éducatifs multilingues, de rapprocher les instruments des conditions effectives d'apprentissage et de renforcer l'équité de la mesure, tout en exigeant un travail plus poussé d'harmonisation, d'adaptation et de vérification linguistique afin de préserver l'équivalence des items entre versions.

Le cycle 2024 enrichit également l'analyse de l'environnement des apprentissages. La composante consacrée aux enseignants, introduite en 2019, est maintenue et renforcée ; les chefs d'établissement demeurent une source centrale d'informations sur les établissements ; les parents ou familles sont intégrés parmi les répondants afin de mieux documenter l'environnement familial, le soutien aux apprentissages et certains déterminants extrascolaires. Une attention plus explicite est par ailleurs accordée à l'inclusion et à la prise en compte de la diversité des élèves dans les procédures d'administration.

Une autre innovation structurante concerne la transformation numérique de la collecte. Pour la première fois dans une évaluation internationale PASEC, les élèves de début de primaire sont évalués au moyen de tablettes et d'une application CAPI (Computer-Assisted Personal Interviewing). Cette évolution modernise la chaîne de collecte, facilite certains contrôles de cohérence et contribue à la sécurisation et à la standardisation des données, tout en maintenant des dispositifs adaptés aux différentes composantes de l'enquête.

5. Lecture synthétique de l'évolution des trois cycles

Dimension	PASEC2014	PASEC2019	PASEC2024
Positionnement	1 ^{er} cycle international standardisé	2 ^e cycle ; consolidation et tendances	3 ^e cycle ; changement d'échelle et modernisation
Pays participants	10	14	16
Niveaux élèves	Début et fin du primaire	Début et fin du primaire	Début primaire, fin primaire et collège / fin de l'éducation de base
Domaines élèves	Langue/lecture et mathématiques	Langue/lecture et mathématiques	Lecture/langue et mathématiques, avec extension au collège
Enquête enseignants	Données contextuelles sur les enseignants	Tests disciplinaires et didactiques + questionnaire enseignant	Composante enseignants renforcée ; volume nettement accru
Familles / parents	Informations familiales principalement via les élèves	Informations familiales principalement via les élèves	Parents/familles intégrés comme répondants
Chefs d'établissement	Questionnaires directeurs	Questionnaire Directeur/École	Chefs d'établissement intégrés à la collecte
Collecte	Principalement papier-crayon ; administration individuelle en début de primaire	Papier-crayon ; cahiers tournants ; procédures QA renforcées	Introduction du CAPI/tablettes en début de primaire ; dispositifs papier et numériques selon les composantes
Nombre et types de langues d'évaluation	3 langues : français, anglais et kirundi	7 langues : français, anglais, arabe et langues nationales africaines	14 langues d'enseignement, combinant langues internationales/officielles, dont l'anglais et le portugais, et langues nationales
Langues nationales prises en compte	Kirundi	Kirundi, malagasy, haoussa et zarma	Élargissement à plusieurs langues nationales, dont kirundi, malagasy, haoussa, zarma, bamanankan, wolof, soninké, sérère, pulaar et mandingue
Innovation et adaptation linguistique	Première prise en compte de contextes linguistiques différenciés dans un cadre international commun	Diversification des versions et formalisation des procédures de traduction, conciliation, vérification et validation	Changement d'échelle multilingue : extension des langues internationales, meilleure prise en compte des langues nationales d'enseignement et renforcement de l'assurance qualité linguistique
Comparabilité temporelle	Création de la référence	Mesure des tendances 2014–2019 pour les pays communs	Continuité des séries du primaire + nouvelles composantes
Innovation dominante	Métrique internationale commune	Évaluation des compétences des enseignants	Collège, familles, numérique, inclusion, élargissement géographique et linguistique
Lien avec l'ODD 4	Contribution à la mesure des apprentissages	Suivi renforcé des apprentissages et de l'équité	Couverture élargie de la cible 4.1.1 jusqu'à la fin du premier cycle du secondaire

En synthèse, PASEC2014 a standardisé et créé la référence internationale ; PASEC2019 a consolidé cette référence, rendu possible l'analyse des tendances et intégré les compétences des enseignants ; PASEC2024 conserve les séries comparables tout en élargissant l'évaluation à de nouvelles populations et à la fin de l'éducation de base, et en modernisant

les modalités de collecte. Cette progression traduit une évolution d'un dispositif centré sur la mesure des acquis des élèves vers un système plus complet sur les conditions de production des apprentissages et les leviers d'amélioration de la qualité de l'éducation.

PASEC

Évaluer les apprentissages pour transformer les politiques éducatives

Créé en 1991, le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) s'est imposé comme l'instrument de référence au service du pilotage des systèmes éducatifs des États et gouvernements membres de la CONFEMEN. Il a pour vocation d'éclairer la décision publique en mesurant l'évolution des performances des systèmes éducatifs, pour orienter l'élaboration, le suivi et l'évaluation des politiques éducatives et, à travers elles, l'amélioration de la qualité de l'éducation.

Son action ne se limite pas à la mesure. Le PASEC renforce les capacités nationales d'évaluation, favorise l'appropriation et l'usage effectif des résultats par les acteurs éducatifs, et soutient l'essor de la recherche en éducation. Par ses évaluations et ses travaux d'analyse, il met à la disposition des décideurs des données fiables et comparables ainsi que des analyses rigoureuses, leviers d'une éducation plus performante, plus équitable et plus efficace.

Le PASEC concourt enfin à la production de connaissances en accompagnant les chercheurs et en mettant à leur disposition ses données et ses cadres d'évaluation. Il nourrit des travaux utiles tant aux systèmes éducatifs qu'à la communauté universitaire.

CONFEMEN – PASEC

Complexe SICAP Point E, Immeuble C
3^e étage, avenue Cheikh Anta Diop
BP 3220 Dakar, Sénégal
Tél. : +221 33 859 29 79 / +221 33 859 29 91
Fax : +221 33 825 17 70
Courriel : pasec@confemen.org
Site web : www.pasec.confemen.org

@CONFEMEN



Avec le soutien de

