

Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles ont-elles un effet sur les performances des élèves de l'enseignement primaire comparativement à ceux évoluant dans des classes « traditionnelles » ? Une revue systématique de la littérature scientifique

Auteur : Peugnieu, Carole

Promoteur(s) : Baye, Ariane

Faculté : Faculté de Psychologie, Logopédie et Sciences de l'Éducation

Diplôme : Master en sciences de l'éducation, à finalité spécialisée en enseignement

Année académique : 2023-2024

URI/URL : <http://hdl.handle.net/2268.2/19999>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles ont-elles un effet sur les performances des élèves de l'enseignement primaire comparativement à ceux évoluant dans des classes « traditionnelles » ?

Une revue systématique de la littérature scientifique.

Mémoire présenté par PEUGNIEU Carole en vue de l'obtention du grade de Master en Sciences de l'Éducation à finalité spécialisée en enseignement

Promotrice : BAYE Ariane

Assistant/chercheur superviseur : DACHET Dylan

Lectrices : FAGNANT Annick et BRICTEUX Sophie

Remerciements

Avant de commencer ce mémoire, je souhaiterais remercier plusieurs personnes :

Merci à ma promotrice Madame Baye et son assisant/chercheur Monsieur Dachet pour leur expertise, leurs précieux conseils, leur aide et leur bienveillance tout au long de la préparation de ce mémoire.

Merci Renaud, Melissa et Hélène pour la relecture et la correction de ce travail.

Merci à mes parents, pour leurs encouragements et pour l'aide apportée au niveau organisationnel.

Merci à mes enfants, Romain et Marion, pour leur patience, leurs encouragements et leur compréhension durant la préparation de ce mémoire mais aussi durant ces années d'études.

Et enfin, merci à mon compagnon sans qui tout ceci n'aurait pas été possible. Merci pour sa motivation, sa disponibilité et pour son soutien sans faille..

Sommaire

1.	Éléments de contexte et de projet	1
2.	Revue de la littérature	2
2.1.	La pratique innovante en éducation	2
2.1.1	Qu'est-ce qu'une pratique ?	2
2.1.2.	L'innovation en éducation	2
2.1.3.	La pratique innovante et la recherche	4
2.1.4.	L'innovation dans le Pacte pour un enseignement d'Excellence	5
2.2.	L'efficacité des pratiques en éducation	6
2.3.	Les formations en Fédération Wallonie Bruxelles (FWB).	9
3.	Méthodologie	12
3.1.	Les pratiques innovantes sélectionnées	12
3.4.1.	L'école du dehors	12
3.4.2.	Les cartes mentales	13
3.4.3.	La classe flexible	13
3.4.4.	L'éveil aux langues	14
3.2.	Les formations en lien avec les pratiques innovantes	15
3.3.	Question de recherche	16
4.	Hypothèse de départ	16
4.1.	Critères d'éligibilité communs aux quatre pratiques innovantes	16
4.2.	Bases de données utilisées	18
4.3.	Équations de recherche	18
4.4.	Sélection des études et extraction des données	18
4.5.	Calculs des amplitudes de l'effet	19
4.6.	Interprétation de l'ampleur de l'effet moyen	20
4.7.	Critères de qualité des études incluses	21
5.	Présentation des résultats	22
4.1.	Résultats et discussion « école du dehors »	22
4.1.1.	Flowchart	22
4.1.2.	Caractéristiques des études incluses	23
4.1.3.	Description générale des études « école du dehors	25
4.1.4.	Discussion « école du dehors »	28
4.1.5.	Conclusion des études pratiquant « l'école du dehors »	30

4.2. Résultats et discussion : Les cartes mentales	31
4.2.1. Flowchart	31
4.2.2. Caractéristiques des études incluses	32
4.2.3. Description générale des études utilisant des cartes mentales	34
4.2.4. Discussion	37
4.2.5. Conclusion des études	39
4.3. Résultats et discussion : La classe flexible	41
4.3.1. Flowchart	41
4.3.2. Caractéristiques des études incluses	42
4.3.3. Description générale des études	43
4.3.4. Discussion	44
4.4. Résultats et discussion : Éveil aux langues	46
4.4.1. Flowchart	46
4.4.2. Caractéristiques des études incluses	46
4.4.3. Description générale des études	50
4.4.4. Discussion	52
4.4.5. Conclusion des études	54
5. Limites et perspectives	56
6. Conclusion	57
7. Bibliographie	58
8. Annexe	66
Annexe 1 Qualité des études	66
Annexe 2 Programmes détaillés des études	73
Annexe 3 Outils de récolte de données des études	78
Annexe 4 Équations de recherche	82

1. Éléments de contexte et de projet

Les résultats médiocres de la Belgique francophone aux enquêtes internationales sont indubitables : l'intérêt pour la lecture est en diminution, le taux de redoublement est extrêmement élevé malgré son inefficacité prouvée, les inégalités liées à l'origine sociale sont trop marquées, les performances des élèves en lecture diminuent, etc. (Portail FW-B, 2019). Assurément, le système scolaire belge et, par conséquent, ses équipes éducatives ont en perspective de nombreuses tâches à accomplir.

L'ensemble des réformes mises en place actuellement par « Le pacte pour un enseignement d'Excellence » a pour objectif de renforcer et d'améliorer l'enseignement actuel. Cependant, les enseignants belges, même s'ils sont contraints de suivre ces changements institutionnels, gardent une certaine autonomie dans le choix de leurs pratiques de classe (Budel, 2020). Pour remédier au manque d'efficacité du système éducatif, une partie d'entre eux ont tendance à se tourner vers des pratiques dites « innovantes », considérées comme économiques par rapport à leur efficience et leur productivité (Charitonidou & Ioannitou, 2012). Ces enseignants cherchent des solutions là où les institutions échouent. Mais ces pratiques améliorent-elles réellement les performances des élèves ? Les croyances de ces enseignants sont-elles fondées ? Sont-ils influencés par des phénomènes de mode ou suivent-ils vraiment la voie de l'efficacité ? La Fédération Wallonie Bruxelles (FW-B) semble valider leurs choix en proposant plusieurs formations sur les pratiques innovantes comme « l'école du dehors », les classes flexibles, le « Mind Mapping », le « sketchnoting » ou encore l'éveil aux langues. En effet, les politiques de formation continuée de la FW-B accompagnent les enseignants tout au long de leur carrière afin d'assurer la continuité de leur formation initiale et permettre l'évolution de leurs pratiques en fonction des élèves d'aujourd'hui. Cependant, se base-t-elle sur des dispositifs véritablement efficaces ? Permettent-ils aux élèves de l'enseignement primaire d'atteindre les résultats escomptés ? Dans ce rapport, nous tenterons de déceler les effets de ces pratiques qui se répandent de plus en plus dans les écoles, afin que les enseignants à la recherche d'innovation puissent baser leurs actions sur des données d'un haut niveau de preuve, la question de recherche étant : « *Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles ont-elles un effet sur les performances des élèves de l'enseignement primaire comparativement à ceux évoluant dans des classes « traditionnelles ? Une revue systématique de la littérature scientifique.*

2. Revue de la littérature

Préalablement, il est élémentaire de tenter de définir quelques notions essentielles à la compréhension des concepts de ce rapport, en lien avec notre question de recherche et la littérature scientifique. Nous présentons, premièrement, des notions de « pratiques innovantes » d'un point de vue universel avant de nous focaliser sur celles dans l'univers de l'éducation. Une fois cette notion bien définie, nous porterons notre intérêt sur l'efficacité d'une pratique dans notre système scolaire et nous terminerons par l'analyse des formations spécifiques proposées aux enseignants de la Belgique francophone.

2.1. La pratique innovante en éducation

2.1.1. Qu'est-ce qu'une pratique ?

Nous allons abondamment parler de pratiques dans ce rapport mais que signifie réellement cette notion ? Une pratique est définie comme l'acquisition d'un assortiment d'habilités chez un individu grâce à un entraînement régulier d'une action (Legendre, 2005, cités par Potvin, 2015). Il n'existe cependant pas de définition unique et il semble difficile de l'expliquer en quelques mots seulement. Globalement, nous pourrions la décrire comme un procédé professionnel exécuté par un spécialiste (Potvin, 2015). Cependant, de multiples facteurs peuvent influencer une pratique : la formation et la personnalité du professionnel, le contexte dans lequel elle est utilisée ou encore l'évolution quotidienne de son développement (Cros & Broussal, 2020). La pratique d'enseignement se déroule généralement en 3 étapes : la planification, les interactions avec les élèves ainsi que les feedbacks et évaluations se déroulant après la période d'enseignement (De Saint André et al., 2013).

2.1.2. L'innovation en éducation

L'innovation en éducation est, quant à elle, une notion qui, au préalable, semble facilement compréhensible et pourrait se traduire par la nouveauté et le changement. Toutefois, si nous analysons les recherches, le sens qui lui est donné peut facilement devenir plus nébuleux et complexe à définir (Potvin, 2015). Le terme innovation est souvent utilisé dans le milieu de la technologie et des sciences. Cependant, il a semblablement sa place dans le monde de l'enseignement. Se contenter de la définir par la nouveauté n'est pas suffisant, même s'il s'agit de sa principale spécificité (Potvin, 2015). Cros et Broussal (2020) ont analysé 500 définitions dans divers domaines et ont pu relever des constantes. Ils associent l'innovation à un

changement réfléchi et dépendant de la motivation de l'innovateur ; ils soulignent que cette énergie motivationnelle est liée à une envie d'amélioration.

L'innovation résulte d'une intention et met en œuvre une ou des actions visant à changer ou modifier *quelque chose* (un état, une situation, une pratique, des méthodes, un fonctionnement), à partir d'un diagnostic d'insuffisance, d'inadaptation ou d'insatisfaction par rapport aux objectifs à atteindre, aux résultats, aux relations de travail. (Cros, 2001, cité par Potvin, 2015, paragr. 4)

Il est primordial que l'objectif premier de l'innovation en éducation soit la réussite de tous les élèves, peu important leurs spécificités, et l'amélioration durable et constante de leur évolution (CSÉ, 2006).

Nous pouvons ensuite distinguer deux types d'innovation : elle peut être soit « incrémentale », son objectif étant d'améliorer un processus progressivement sans en modifier complètement les bases ; soit « majeure », en induisant une transformation de la structure ou de l'organisation de l'institution. Cette innovation peut cependant être difficile à adopter par l'enseignant, car elle est souvent imposée par le système alors que le changement doit idéalement venir des acteurs (Cros & Broussal, 2020).

Si l'enseignant souhaite mettre en place une innovation pédagogique, il est préférable qu'il ait un minimum d'autonomie ainsi qu'une certaine confiance en lui afin de parvenir à ses fins. Comme le soulignent Charitonidou & Loannitou (2012) dans leur recherche analysant les caractéristiques des enseignants qui influencent la mise en place d'innovations pédagogiques, l'enseignant ne doit pas avoir peur de s'engager ; il doit avoir pour principal objectif les progrès de tous ses élèves. Le fait d'entretenir une bonne relation avec eux, de maîtriser leurs niveaux et la matière à enseigner est, par ailleurs, nécessaire pour implémenter l'innovation avec succès. L'accès à des outils et du matériel appropriés et une formation initiale (ou continue) adaptée sont également des atouts pour oser se lancer. Ces auteures ajoutent que les enseignants doivent accepter de sortir des sentiers battus afin de pouvoir « construire le nouveau » (p.59).

D'après Potvin (2015), l'innovation en éducation se définit essentiellement par ces quatre concepts :

- « Une nouveauté » qu'il dissocie de la novation ou de l'invention, car celle-ci est une conception nouvelle n'ayant aucune autre présence ailleurs tandis que l'innovation (qu'elle soit incrémentale ou majeure) se caractérise par la nouveauté en lien avec un contexte existant.

- « Un produit » : une pratique, un concept, un processus, etc., les éléments auxquels nous prêtons cette innovation ;
- « Un changement et une amélioration » : l'innovation a pour objectif de tenter de résoudre une situation problématique en améliorant l'existant. Cette démarche doit idéalement provenir de son initiateur et, par conséquent, être un acte volontaire.
- « Une durabilité et une transférabilité » : l'innovation ne doit pas rester un acte isolé, localisé et temporaire. Pour être efficace, elle doit se répandre et être exploitée par d'autres acteurs de l'enseignement, au-delà de celui qui l'a créée, en subissant des adaptations et des changements si nécessaire (2015, Ce qui caractérise l'innovation en éducation).

En outre, il est fondamental de mettre en évidence que certaines pratiques dites « innovantes » correspondent à des pratiques ayant plusieurs décennies. Tricot (2017), dans son livre « L'innovation pédagogique », met en exergue que l'innovation pédagogique a tendance à utiliser d'anciennes méthodes en les faisant passer pour des nouvelles. Ce point fera l'objet de vigilance lors de la sélection des pratiques analysées dans ce rapport.

Dans le même ordre d'idées, il est notable de dissocier les pratiques innovantes des pédagogies alternatives. Ces dernières s'opposent aux pédagogies dites « traditionnelles » (Wagnon, 2018) et se qualifient de pédagogies actives plaçant l'enfant au centre de ses apprentissages tout en respectant ses besoins, ses centres d'intérêt et sa liberté (Fournier, 2016). En d'autres mots, elles centrent majoritairement leurs attentions sur les besoins individuels des enfants. Ces critères spécifiques ne font pas nécessairement partie de ceux proposés dans les pratiques innovantes. De plus, les pédagogies alternatives se revendiquent implicitement ou explicitement comme appartenant au courant d'Éducation Nouvelle. Ce courant, influencé par Freinet, Decroly, Montessori, etc., n'est pas nouveau, malgré son nom, puisqu'il trouve ses racines avant le début du 20^e siècle. L'innovation, une fois établie depuis un moment, devient normale et n'est par conséquent plus de l'innovation (De Ketele, 2002, cité par Cocton, 2019). Cependant, il n'est, en rien, exclu que les praticiens des pédagogies alternatives utilisent épisodiquement des pratiques innovantes. Il est simplement utile de ne pas les confondre, l'une n'allant pas nécessairement avec l'autre.

2.1.3. La pratique innovante et la recherche

Comme abordé dans le préambule de ce rapport, les enseignants ont tendance à se tourner vers des pratiques innovantes afin de pallier le manque d'efficacité de notre système éducatif. En effet, l'innovation en éducation survient généralement lorsque l'enseignant rencontre des difficultés sur le terrain de manière récurrente et qu'il s'essaie à de nouvelles façons d'enseigner. Cette pratique innovante peut devenir une pratique « exemplaire » lorsque

l'innovateur collabore avec un chercheur afin d'élaborer « le processus de recherche et d'évaluation » (Potvin, 2015, De la pratique à la recherche). Toujours selon cet auteur s'inspirant de Brodin (s.d.), l'innovation possède un processus en trois étapes et en trois temps : « l'expérimentation pilote, les adaptations aux réalités du terrain et l'institutionnalisation ou sa généralisation » (Les étapes dans le processus d'innovation).

Inversement, l'innovation peut également provenir des résultats de recherches scientifiques (CSÉ, 2006). C'est le cas lorsqu'un enseignant modifie ses pratiques en utilisant la recherche.

Dans les deux situations, les échanges entre la pratique et la recherche semblent stimuler l'innovation en éducation (Potvin, 2015). Cet auteur précise que l'innovation vient d'une remise en question et d'une prise d'initiative mais, énonce également le caractère « créatif et imaginatif » que cela implique. Il conseille de se baser sur les connaissances et les expérimentations des autres professionnels ainsi que sur les théories proposées sur le sujet. (Cros, 2002 ; CSÉ, 2006 cités par Potvin, 2015).

2.1.4. L'innovation dans le Pacte pour un enseignement d'Excellence

Le Pacte pour un enseignement d'Excellence, cet ensemble de réformes systémiques de grande envergure débuté en 2015 dans l'enseignement belge francophone visant à améliorer la qualité de l'enseignement (Pacte pour un Enseignement d'excellence, 2023), se dit vouloir, lui aussi, être innovant de diverses manières, comme le précise la Fédération Wallonie Bruxelles (FWB) sur son site :

Il introduit une démarche innovante sous différents aspects :

- « [...] Il propose avant tout de remettre la pédagogie au centre du processus (et non les systèmes) et de déployer l'autonomie et l'innovation pédagogique des acteurs.
- [...] Il refuse les démarches dirigistes et opte pour un large processus participatif innovant, impliquant aussi les acteurs hors enseignement.
- [...] Il propose de développer une culture d'évaluation, aborde le défi de la modernité, de la créativité et de la transition numérique » (Pacte pour un Enseignement d'Excellence, 2015, Adoption par le gouvernement de la FWB).

Afin de soutenir les écoles étant identifiées comme étant en écart de performances, le Pacte dit vouloir mettre, à la disposition des enseignants, des programmes innovants ayant fait leurs preuves.

2.2. L'efficacité des pratiques en éducation

Comme cité précédemment, les enquêtes internationales telles que le Programme International pour le Suivi des Acquis des élèves (PISA) ont pu mettre en évidence de nombreuses anomalies dans le fonctionnement du système belge francophone. Les innombrables réformes éducatives mises en place antérieurement semblent vraisemblablement inefficaces. Elles se veulent parfois alléchantes, parfois engageantes, parfois idéalistes, etc. Cependant, elles sont rarement basées sur des pratiques ayant une efficacité prouvée (Draelants & Dupriez, 2015).

Selon de nombreux auteurs (Baye, 2022 ; Baye & Bluge, 2016 ; Bissonnette, 2020 ; Camus, 2019 ; Davies, 1999 ; Dachet & Baye, 2020 ; Slavin, 2019), l'éducation basée sur les preuves constituerait une clé pour parvenir enfin à des résultats. « Faire des preuves une base pour l'adoption de programmes placerait l'éducation dans un cercle vertueux d'innovation, d'évaluation et d'amélioration progressive comme celui qui a transformé des domaines tels que la médecine, l'agriculture et la technologie » (Slavin, 2020, page 22).

L'éducation basée sur des preuves «Evidence Based Education (EBE)» ou «approche basée sur des interventions éducatives dont l'efficacité a été prouvée» (Baye, 2018, citée par Bissonnette, 2020, p. 2) est apparue dans les années 1990 aux États-Unis et au Royaume-Uni. Cette approche présente aux enseignants et autres acteurs de l'éducation des dispositifs prouvés scientifiquement suivant une méthodologie stricte (Dachet & Baye, 2021). D'après Bissonnette (2020), si les dispositifs utilisés ont des hauts niveaux de preuves, des « données probantes », des changements et des améliorations sont concevables en éducation.

Ces pratiques sont validées par des preuves scientifiques, contrairement à celles basées sur les traditions et les croyances ayant un caractère souvent subjectif (La Roche, 2008, cité par Bissonnette, 2020). D'après certains auteurs, la précision mobilisée pour prouver l'efficacité de pratiques n'a aucun lien avec les intuitions ni les certitudes personnelles (De Chambrier & Dierendonck, 2022). Propos confirmés par Carnine (2000, cité par Bissonnette, 2020), qui désigne de manière plus nuancée l'enseignement comme une profession immature donnant plus d'importance aux croyances qu'aux résultats minutieusement contrôlés par les scientifiques. Il revendique l'importance de se fier à des données objectives afin de définir l'efficacité d'un dispositif.

Slavin (2013, cité par Baye & Bluge, 2016) dénonce quant à lui les réformes éducatives dirigées par des pressions extérieures, ou par des effets de mode comme « l'introduction de l'approche par compétence » (page 6) qui est une réforme de grande étendue n'ayant pourtant

pas d'appui scientifique (Castonguay & Gauthier, 2012 ; Hattie, 2009 ; Lafontaine, 2016 cités par Baye & Bluge, 2016). Il prône la mise en place de réformes basées sur des pratiques jugées efficaces en passant par des études expérimentales (Baye & Bluge, 2016). Dachet et Baye (2021) confirment le manque d'efficacité prouvée « des socles de compétences » en s'appuyant sur des revues systématiques comme celle de Morcke et al. (2013) ou une revue de la littérature de Lassnigg (2015) et précisent que même si un équilibre est prévu dans le Pacte avec la réinsertion des savoirs, l'approche par compétences reste d'actualité malgré son efficacité non établie.

Pourtant, chaque année, des recherches en éducation sont réalisées par les Universités et les Hautes Écoles de la Fédération Wallonie Bruxelles afin de conseiller les enseignants ou autres professionnels de l'éducation. Elles mettraient à leur disposition les recherches effectuées et collaboreraient avec l'Institut inter réseaux de la Formation Professionnelle Continue (IFPC) afin que leurs formations puissent être en adéquation (Enseignement.be, n.d.). Cependant, plusieurs auteurs critiquent la formation des enseignants comportant, selon eux, de nombreuses lacunes notamment celle de ne pratiquement pas avoir recours aux interventions fondées sur des données probantes (Carnine, 2000, Sauvé, 2012, Slavin 2002 cités par Bisonnette, 2021).

Il est notable de préciser que l'EBE ne fait pas l'unanimité. En effet, certains acteurs de l'enseignement, reprochent aux chercheurs de ne pas rendre suffisamment exploitables les résultats de leurs études ou de ne valider que les résultats provenant d'études expérimentales ou quasi expérimentales ; ce qui implique selon eux « une centration excessive sur les objectifs ou les produits mesurables » (Lessard, 2007, cité par Baye & Bluge, 2016, p.7). Les experts pensent, par ailleurs, que cette méthodologie est une des seules capables de produire des données fiables. En effet, afin que les recherches puissent amener des données probantes, elles doivent remplir plusieurs critères strictement scientifiques :

- Formuler des questions de recherche claires et pertinentes qui peuvent être l'objet d'une investigation empirique ;
- Se fonder sur une ou plusieurs théories reconnues et pertinentes ;
- Recourir à des méthodes qui permettent d'investiguer directement sur l'objet de la question de recherche ;
- Faire preuve d'une chaîne de raisonnements cohérents et explicites théoriquement et empiriquement fondés ;
- Être reproductible et pouvoir fournir des résultats généralisables ;
- Faire l'objet de publications substantielles et complètes permettant et encourageant la critique et l'évaluation de la part de la discipline et de la profession.

(Larose, Couturier, Bédard et al., 2011, p. 36)

Cependant, il existe une quantité invraisemblable d'articles scientifiques et il est, par conséquent, parfois difficile d'identifier les données fournies sur un sujet en particulier. Les synthèses de données probantes permettent d'obtenir une vue d'ensemble sur les recherches existantes dans le domaine. Plusieurs méthodes de synthèse permettent d'amener un haut niveau de preuves comme, par exemple, la revue systématique. Elle permet d'examiner, à l'aide d'un protocole précis, tous les résultats de la recherche sur un thème bien défini de manière structurée et minutieuse. Elle cherche à démontrer si des preuves sont suffisamment robustes dans la littérature scientifique pour répondre à une question de recherche. Différentes étapes doivent être respectées :

- « La formulation de la question de recherche ;
- la définition des critères d'inclusion [et d'exclusion] ;
- l'identification des études ;
- l'extraction des données et la synthèse des données »

(Gough, 2007, cité par Sacré, Lafontaine & Toczec, 2021, p. 5).

Les revues systématiques peuvent mettre en évidence des problèmes de fiabilité de données, permettre un gain de temps considérable lors de l'analyse et du tri d'informations sur un sujet recherché et d'orienter des décisions au moment de choisir une pratique à mettre en place.

Dans la recherche scientifique, les revues systématiques sont primordiales pour la progression des connaissances. Elles permettent d'observer certains manques et de faire apparaître de nouvelles recherches (Young & Eldermire, 2017).

Quoi qu'il en soit, l'efficacité est au cœur des préoccupations et les systèmes éducatifs anglais et francophones s'intéressent progressivement et amplement à l'éducation basée sur les preuves et à la gestion par les résultats. Au Royaume-Uni, les experts qui conseillent le gouvernement ont publié un guide pédagogique, venant des résultats des recherches scientifiques, reprenant des prototypes d'exercices et des recommandations à destination des intervenants éducatifs (L'éducation fondée sur des preuves: définition et enjeux , 2022).

Concernant, plus précisément, la Fédération Wallonie Bruxelles, dans le Pacte pour un Enseignement d'Excellence, il est aussi recommandé de tenir compte des résultats des recherches scientifiques plus largement ainsi que des données probantes (Baye, 2018, citée par Bissonnette, 2020). En effet, un rapport réalisé par des experts « L'éducation fondée sur des preuves » (Baye & Bluge, 2016) a été mandaté dans le cadre du Pacte. Il propose cinq pistes

permettant l'implémentation de l'EBE en Belgique francophone en tenant compte de ses forces et ses faiblesses :

Piste 1. Éviter le piège de la technocratie [...]

Piste 2. Approfondir la question de l'articulation des savoirs expérimentiels et des savoirs de la recherche [...]

Piste 3. Prioriser les domaines d'intervention [...]

Piste 4. Définir la structure et les standards de recension [...]

Piste 5. Impliquer les acteurs de l'école dans les prises de décisions. [...]

(Baye & Bluge, 2016, pp. 9-15)

Grâce à ce rapport, l'EBE a été reconnu, dans l'avis n°3 réalisé par le groupe central gérant la réforme du Pacte, comme un outil permettant de donner aux acteurs de l'enseignement des pistes concrètes éprouvées pour enrichir leurs pratiques (Dachet et Baye, 2021). L'avis n°3 a été adopté par le gouvernement en 2017 et propose une série de mesures à implémenter pour améliorer l'enseignement.

La question est maintenant de savoir si les différents dispositifs mis en place pour les enseignants, et plus précisément dans les formations en cours de carrière proposées par notre système éducatif, se basent réellement sur ces recherches scientifiques. Permettent-ils aux équipes éducatives d'utiliser des pratiques reconnues comme efficaces et par la même occasion de lutter activement contre l'échec scolaire, le décrochage ou encore le redoublement faisant malheureusement trop souvent parties des modes de fonctionnement de la Belgique francophone ?

2.3. Les formations en Fédération Wallonie Bruxelles (FWB)

Afin de répondre à notre question de recherche, nous passerons en revue ultérieurement les différentes formations proposées par la FWB à ses enseignants du niveau primaire. Cependant, il semble opportun de décrire ou de rappeler brièvement, au préalable, le fonctionnement spécifique des formations du système scolaire belge actuellement en plein changement par l'intermédiaire du Pacte pour un enseignement d'Excellence.

Après plusieurs reports et débats, de nombreuses adaptations ont eu lieu très récemment (septembre 2023) au niveau de la formation initiale des enseignants, comme notamment l'allongement du cursus de 3 ans à un cursus de 4 ans, dont une partie au niveau universitaire.

Les enseignants seront, idéalement, mieux formés pour faire face à l'hétérogénéité et aux difficultés d'apprentissage de leurs étudiants et mieux préparés aux réalités du terrain afin de lutter contre les faiblesses de notre système mises en exergue dans les enquêtes internationales mentionnées antérieurement.

Étant donné ces nombreux changements en cours, nous centrerons les recherches de ce rapport exclusivement sur les formations « en cours de carrière » au niveau des enseignants primaires. Ces formations subissent actuellement moins de changements au sein du Pacte et permettent une analyse basée sur l'existant. Depuis 2002, la majorité des pays européens, dont la Belgique, ont constitué pour leurs enseignants différentes formes de « formations en cours de carrière » (Laruelle, 2009, p.5). Pour l'OCDE, même si la formation initiale doit être de haute qualité, elle n'est pas suffisante pour permettre aux enseignants d'évoluer efficacement durant toute leur carrière. L'objectif de la formation continue est de tenter de répondre aux besoins de chacun au fil de temps et de permettre d'acquérir de nouvelles compétences tout en étant en adéquation avec la formation initialement reçue (Laruelle, 2009). D'après la circulaire 8742 de la FWB (2022), ces formations continues se présentent selon deux modèles :

- Les formations obligatoires répondant à des besoins collectifs de l'école : elles passent dorénavant de 6 demi-journées à 8 à 12 demi-journées par an en fonction des nécessités. Les besoins doivent être déterminés par l'équipe éducative et faire partie du plan de pilotage de l'établissement.

- Les formations volontaires répondant à des besoins personnels : elles peuvent atteindre 10 demi-journées maximum par an et doivent être validées par la direction. Si celle-ci refuse, elle doit pouvoir en justifier les raisons. Par contre, si la formation est acceptée, l'enseignant est par la suite obligé de s'y rendre, excepté en cas de force majeure (Enseignement.be, Circulaires 8742, 2022).

La formation professionnelle en cours de carrière (ou continue) au niveau fondamental est organisée par les Fédérations de Pouvoirs Organisateurs et le Pouvoir Organisateur Wallonie-Bruxelles Enseignement (WBE) selon deux niveaux :

- En inter-réseaux par l'Institut inter-réseaux de la Formation Professionnelle Continue (IFPC). Il est « l'organisme de référence de la Fédération Wallonie-Bruxelles pour les formations en cours de carrière, initiale, d'accession et d'insertion professionnelle à destination des membres du personnel de l'enseignement » (Institut inter-réseaux de la Formation professionnelle, s. d.). Ils proposent des formations à plusieurs pratiques innovantes.

- Au niveau de chaque réseau : Le Conseil de l'Enseignement des Communes et des Provinces (CECP) représente le réseau officiel subventionné et le Secrétariat Général de

l'Enseignement Catholique (SeGEC) est le représentant et le coordinateur du réseau libre catholique en Communauté française et germanophone. Ces deux associations sans but lucratif (asbl) ont pour objectif d'aider les communes et les provinces à accomplir leurs fonctions d'éducation et d'enseignement. Elles proposent, entre autres, des formations continues sur des pratiques innovantes pour les enseignants du fondamental. Ces formations prennent en considération les caractéristiques du réseau et les besoins d'un pouvoir organisateur ou d'une école (Enseignement.be, s.d.).

Hormis ces formations obligatoires et volontaires organisées durant les périodes scolaires, les enseignants ont bien entendu le loisir et la liberté de se former en dehors des périodes scolaires sur les sujets, concepts, pratiques, pédagogies... qui les passionnent, sans en avvertir spécialement leur direction. De nombreuses associations sans but lucratif proposent des offres de formations autant diversifiées sur les thèmes proposés que sur leur fiabilité. En outre, contrairement aux formations proposées par la FWB, elles sont rarement gratuites et peuvent parfois atteindre des tarifs dispendieux.

Pour un souci de clarté et de méthodologie, nous centrons, dans ce rapport, nos recherches uniquement sur des pratiques innovantes approuvées par la Fédération Wallonie-Bruxelles et proposées aux enseignants dans ses organismes de formations. Cependant, au préalable, il semble nécessaire de décrire ces différentes pratiques afin de permettre aux lecteurs d'avoir une vision précise des concepts utilisés dans la suite de ce rapport.

3. Méthodologie

3.1. Les pratiques innovantes sélectionnées dans les formations en cours de carrière de la Fédération Wallonie-Bruxelles

Nous avons analysé toutes les formations en cours de carrière proposées par la Fédération Wallonie-Bruxelles via ses différents organismes et nous avons sélectionné les plus répandues n'ayant pas encore fait leurs preuves dans l'enseignement primaire ordinaire. A titre d'exemple, nous avons exclu les formations sur l'enseignement explicite (Boucher & Bouffard, 2015) et le programme P.A.R.L.E.R. (Zorman et al., 2015) qui ont déjà démontré leur efficacité. Par conséquent, nous avons retenu quatre pratiques innovantes : (1) l'école du dehors ; (2) les cartes mentales ; (3) la classe flexible ; (4) l'éveil aux langues.

3.1.1. L'école du dehors

Depuis les années 1950, l'école du dehors se répand progressivement. Initialement, cette pratique s'est développée en Europe du Nord, puis plus récemment, elle a gagné en popularité dans de nombreux autres pays développés. Depuis une dizaine d'années, l'école du dehors connaît un développement important en Belgique et les activités en lien avec ce programme sont très variables. Si un des objectifs de cette approche est bien entendu la connectivité avec la nature, d'autres missions sont poursuivies comme le questionnement et les hypothèses, le décroisement des matières scolaires ou encore le ressenti corporel et la découverte (collectif Tous Dehors, s.d.). Ce phénomène a pris encore plus d'ampleur suite au premier confinement durant l'année 2020. Les enseignants ont trouvé, dans cette pratique, un moyen efficace permettant de respecter les normes sanitaires en vigueur.

L'école du dehors propose aux enseignants et aux enfants de sortir des bâtiments et de réaliser des apprentissages à l'extérieur, idéalement dans l'environnement proche de l'école. Elle regroupe des pratiques diverses, éducatives et pédagogiques, en lien avec les référentiels scolaires, dans le but de développer des savoirs, des savoir-faire, et des savoir-être (Spw, 2022).

D'après Acheroy et ses collègues (2020) : « une pratique régulière de sorties dehors est une dynamique riche et unique pour répondre à une série d'enjeux liés aux apprentissages, au développement de l'enfant et à la connaissance de son milieu de vie, ainsi que, plus globalement, à l'éveil d'une conscience citoyenne, dans l'actuel contexte écologique pour le moins préoccupant » (p. 7). Cependant, il peut être plus difficile pour les enfants de canaliser leur attention, de se centrer sur des objectifs, de ne pas se laisser distraire par les nombreux distracteurs présents dans les espaces extérieurs. Des auteurs soulignent les réticences de certains enseignants comme Partoune (cité dans Dehors : la Nature Pour Apprendre, 2013) : des

enseignants peuvent considérer qu'enseigner en plein air est plus risqué. En effet, l'environnement extérieur est dynamique et imprévisible ; les interactions sociales sont différentes, le corps est plus libre et il faut être conscient des dangers potentiels liés à cet environnement. Dans ce rapport, nous tenterons de vérifier l'efficacité de cette pratique en plein essor.

3.1.2. Les cartes mentales

Même si le concept de représentation visuelle remonte à l'antiquité, la carte mentale telle que nous la connaissons aujourd'hui a été développée par le psychologue anglais Tony Buzan dans les années 70 (Dambreville, 2014).

Les cartes mentales (mind maps en anglais) servent à résumer des informations visuellement, en réalisant des liens à l'aide de mots-clés ou de symboles en partant d'un mot central. Les idées ou concepts moins importants sont présentés à travers des branches de plus en plus éloignées du centre. Cette technique a pour objectif une concentration plus ciblée sur le principal de l'information reçue, une mémorisation plus efficace et l'obtention d'une meilleure compréhension et structure (Initiation au mind mapping, 2023). Cette pratique peut se faire à l'aide du papier crayon ou via des logiciels informatiques prévus à cet effet et peut s'étendre à l'infini en fonction des besoins (Yang, 2017).

Certains auteurs soutiennent que les cartes mentales « permettent aux étudiants d'avoir une meilleure compréhension conceptuelle des principes de base lorsqu'il y a beaucoup d'informations, et qu'elles favorisent l'apprentissage actif pour améliorer leur mémoire, en particulier lors de la récupération d'informations écrites » (Aljaser, 2017 ; Nodoushan & Maibodi, 2017; Kaveh, & Rassaei, 2016; Tarkashv, 2015; Carol, 2015; Toi, 2009 cités par Alwattar & Al-Balhan, 2018, p. 116). Nous tenterons de vérifier ces propos dans cette revue.

Le sketchnoting fonctionne avec les mêmes principes que le mind mapping mais dans ce cas, les dessins ou les graphiques sont préconisés même si les mots-clés sont permis. Le but est également d'organiser les informations reçues de manière visuelle afin de les mémoriser à long terme. Sketch signifie « croquis » et noting veut dire « notes » (Jambon, 2022).

3.1.3. La classe flexible

Dans le monde entier, une classe traditionnelle est généralement un espace fermé où des bancs sont alignés souvent par rangées face à un enseignant et à un tableau (noir ou interactif). La classe flexible, quant à elle, aborde l'environnement de la classe de manière beaucoup moins traditionaliste : des tables sont regroupées pour favoriser le travail collaboratif, une bibliothèque et des ordinateurs en nombre plus importants sont parfois à disposition des élèves afin de faciliter les activités de recherches et de lecture, des espaces dans le couloir ou dans un

local à proximité peuvent également être prévus pour les élèves en optimisant de la sorte les surfaces disponibles (Maulini & Capitanescu, 2020).

La popularité et l'augmentation des assises dynamiques et des aménagements flexibles dans les classes est indéniable depuis quelques années. La classe flexible permettrait de lutter contre la sédentarité rencontrée par les enfants de l'enseignement traditionnel. L'espace physique peut améliorer l'apprentissage des élèves et la configuration des espaces d'apprentissage est fondamentale (Brooks, 2011). Mais ce dispositif a-t-il été validé par la littérature scientifique ?

3.1.4. L'éveil aux langues

Les activités connues sous le nom « d'éveil aux langues » ont commencé à émerger en Grande Bretagne dans les années 80. D'après De Pietro & Matthey (2001), leur but était de pallier trois grands problèmes :

« Les difficultés d'intégration et la fréquence des échecs scolaires des élèves issus de la migration, les difficultés des élèves anglophones dans l'apprentissage de langues étrangères et *last but not least*, les difficultés en anglais, attribués au niveau trop limités de leurs capacités méta langagières »
(p. 33).

Dans les années 90, ce concept a émergé dans les pays francophones par Dabène et Perregeaux (cités par De Pietro & Matthey, 2001). En Belgique, l'éveil aux langues ambitionne l'ouverture aux langues étrangères, aux sons et aux rythmes particuliers de celles-ci mais également à sensibiliser les enfants aux autres cultures, aux différences, à l'ouverture d'esprit pour une société plus tolérante (Enseignement.be, s.d.).

Avec l'arrivée du tronc commun, un élargissement de l'apprentissage des langues étrangères est maintenant prévu dès la 3^{ème} primaire et l'enseignement d'une seconde langue étrangère sera bientôt obligatoire dès la 2^{ème} secondaire dans toutes les écoles de la Fédération Wallonie-Bruxelles. Depuis septembre 2020, les titulaires de la 1^{ère} maternelle à la 2^{ème} primaire doivent pratiquer l'éveil aux langues à raison d'une heure par semaine (Enseignement.be, s.d.).

Ces pratiques seront analysées ultérieurement afin de permettre aux enseignants de se baser sur des pratiques ayant prouvé leur efficacité.

3.2. Les formations en lien avec les pratiques innovantes

Chaque pratique a été sélectionnée suite à l'analyse complète de toutes les formations de la FW-B destinées aux enseignants de l'enseignement primaire ordinaire :

Pratiques innovantes sélectionnées.	Intitulés des formations proposées en lien avec la pratique innovante	Organismes proposant cette formation
L'école du dehors	<i>À la découverte de l'environnement proche de l'école : pistes pour oser sortir</i>	CECP
	<i>Interdisciplinarité dans les apprentissages par le projet de l'école du dehors</i>	CECP
	<i>Engagement, curiosité et envie d'apprendre pour tous en lançant un projet d'Ecole du dehors en maternelles, P1 et P2</i>	IFPC
Les cartes mentales (le mind mapping et le sketchnoting)	<i>Structurer, mémoriser et appliquer ses démarches mentales à l'aide du Mind Mapping</i>	IFPC
	<i>Le sketchnoting pour créer des référentiels attrayants.</i>	SeGEC
La classe flexible	<i>L'apprentissage en mouvement(s).</i>	SeGEC
	<i>La classe flexible disciplinaire ou pluridisciplinaire.</i>	IFPC
L'éveil aux langues	<i>Éveil aux langues : intégrer le plurilinguisme dans les apprentissages via la littérature jeunesse.</i>	CECP
	<i>L'éveil aux langues, un jeu d'enfant. Susciter la curiosité, explorer d'autres univers, favoriser le bien-être en classe, partager des découvertes</i>	IFPC
	<i>Journée d'étude autour de l'Eveil aux Langues</i>	IFPC

3.3. Question de recherche

Pour mener à bien cette synthèse de la littérature, la méthode de revue systématique est utilisée. Comme explicité précédemment, celle-ci doit contenir plusieurs étapes essentielles à sa validité. Avant toute chose, il est primordial de vérifier qu’aucune autre revue systématique récente n’a pas déjà été effectuée sur le sujet ; si c’est le cas, il est essentiel de justifier l’utilité d’en élaborer une nouvelle. Ensuite, il est indispensable que la question de recherche soit clairement définie.

Le canevas *Population- Intervention-Comparaison-Outcomes-Study* (PICOS ; Centre for Reviews and Dissemination, 2008) a été utilisé afin de formuler la question de recherche de ce travail :

P	Population	Élèves de l’enseignement primaire ordinaire
I	Interventions	Les pratiques innovantes utilisées en vue d’améliorer les performances scolaires des élèves
C	Comparaison	Au minimum un groupe d’enfants recevant toute autre intervention ou rien
O	Outcomes	Au minimum une évaluation des performances scolaires
S	Study design	Expérimentales et quasi-expérimentales (Pour l’éveil aux langues : Longitudinales et exploratoires autorisées)

Par conséquent, cette recherche tentera d’apporter des éléments de réponse à la question : « *Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles ont-elles un effet sur les performances des élèves de l’enseignement primaire comparativement à ceux évoluant dans des classes « traditionnelles » ?* »

3.4. Hypothèse de départ

En réponse à la revue de littérature qui a été menée, une hypothèse peut être énoncée dans cette recherche :

→ Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles reposent sur des études ayant démontré des effets sur les performances des élèves de l’enseignement primaire ordinaire.

3.5. Critères d’éligibilité communs aux quatre pratiques innovantes

Conformément à la question de recherche, plusieurs critères d’inclusion et d’exclusion ont été identifiés en suivant le canevas PICOS (Centre for Reviews and Dissemination, 2008) :

	Inclusion	Exclusion
Population	Ce travail se centre sur les enfants de 6 à 12 ans fréquentant l'école primaire dans l'enseignement ordinaire.	Les études se basant sur des enfants présentant des troubles d'apprentissage n'ont pas été retenues.
	Les études devaient présenter un échantillon avec des participants répondant à nos critères. Si l'échantillon était plus grand, il fallait une analyse séparée des différents groupes (par exemple, une étude utilisant des participants ayant des troubles d'apprentissage peut être incluse si elle analyse les résultats des enfants sans trouble d'apprentissage séparément).	Les études s'intéressant aux élèves de l'enseignement secondaire ou supérieur ainsi qu'aux enseignants peu importe leur niveau d'enseignement ont été exclues également.
	Le genre et l'ethnie ne constituaient pas un critère d'exclusion.	Les études présentant un nombre de participants inférieur à 20 ont été exclues car plus l'échantillon est petit, plus il y a de chances de ne pas détecter un effet (Sham, 2013) ou au contraire d'avoir des effets extrêmes.
Intervention	L'intervention proposée devait être un dispositif mettant en place une pratique innovante sélectionnée visant à améliorer les performances des élèves.	Les études utilisant des pratiques innovantes autres que celles proposées par la FW-B étaient exclues.
	Le type d'intervention, la combinaison de plusieurs dispositifs, le nombre de séances et le lieu ne constituaient pas un critère d'exclusion.	
Comparaison	Un groupe contrôle devait exister dans l'étude ou un plan à mesures répétées.	

Outcomes	L'étude devait évaluer l'effet de l'intervention sur les performances scolaires, peu importe la discipline.	Les études évaluant le bien-être, la motivation ou la perception des élèves ont été exclues.
Study design	Le type d'études recherchées était des études expérimentales ou quasi-expérimentales car elles permettent d'apporter un haut niveau de preuve pour l'évaluation d'une pratique.	Les études n'étant pas des études expérimentales ou quasi-expérimentales ont été exclues.

3.6. Bases de données utilisées

Nous avons utilisé deux bases de données pour réaliser cette revue systématique :

La première est ERIC (Education Resources Information Center). Cette base de données est parrainée par l'Institute of Education Sciences du Ministère de l'Enseignement américain et est un outil « de référence » pour les chercheurs en Sciences de l'Éducation ou tout autre personne s'intéressant aux Sciences de l'Éducation ou l'Enseignement (*ERIC | EBSCO*, s. d.).

La deuxième base de données, SCOPUS, permet d'obtenir des références bibliographiques, des résumés et des citations dans des domaines multidisciplinaires. Elle propose la littérature scientifique évaluée par les pairs avec des mises à jour quotidiennes.

3.7. Équations de recherche

Chaque pratique innovante a fait l'objet d'une équation de recherche spécifique dans les bases de données avant d'être exportée dans quatre sessions indépendantes sur le site Rayyan. Ce site a été développé par des chercheurs afin de guider la sélection d'articles scientifiques en (1) appliquant des étiquettes avec texte libre, permettant d'attribuer des commentaires descriptifs sur les études et (2) en facilitant le filtrage en proposant de justifier le choix de l'exclusion (<https://rayyan.ai>). Les quatre équations de recherche se trouvent en annexe 4.

3.8. Sélection des études et extraction des données

Cette tâche est partagée en plusieurs étapes progressive. Premièrement, les doublons sont automatiquement supprimés et la sélection des articles pertinents se réalise par la lecture des résumés puis des études en entier si celles-ci sont disponibles en ligne. Grâce au réseau « ResearchGate », nous avons demandé l'accès directement aux auteurs pour certains articles indisponibles. Six d'entre eux ont répondu favorablement et nous ont permis d'avoir toutes les données nécessaires à la sélection. Ensuite, lors du tri et de l'exclusion des études, nous avons justifié la raison de celle-ci. Par exemple, nous avons rencontré diverses études se préoccupant

de l'avis des enseignants. Même si le sujet abordé correspondait parfaitement à une des pratiques innovantes sélectionnées, nous avons exclu ces études en précisant qu'il s'agissait d'une mauvaise population. La dernière étape consistait à collecter les données nécessaires en lisant entièrement les études incluses dans notre revue.

Nous avons extrait les données suivantes des différents articles en suivant les recommandations de Sacré, Lafontaine & Toczek (2021) :

- Les données générales de l'article : les auteurs, le titre de l'article, la date de publication et le pays.
- Les données sur les participants : leur(s) âge(s), le nombre de participants.
- Les données sur le type d'études utilisées.
- Les données sur la comparaison : la présence de pré test, post test et le nombre de groupes expérimentaux et de contrôle.
- Les données sur les résultats : les variables indépendantes et dépendantes, les tailles de l'effet lorsqu'elles étaient directement disponibles ou, à défaut, les moyennes et les écarts type afin de calculer les ampleurs de l'effet nécessaires.

3.9. Calculs des ampleurs de l'effet

Les ampleurs de l'effet, lorsqu'elles n'étaient pas directement fournies par les auteurs, ont été calculées à l'aide de l'outil en ligne «Psychometrica». Ce dernier propose plusieurs calculatrices permettant de chiffrer une ampleur de l'effet pour les différentes configurations d'études et en fonction des données disponibles. La taille de l'effet est, selon Kamper (2019) : « un moyen de mesurer la force de la relation statistique entre deux variables. Il s'agit d'une mesure standardisée qui est utilisée pour indiquer l'ampleur de l'effet d'une intervention ou d'un traitement » (p.559). Une des mesures de l'effet la plus répandue en sciences de l'éducation est le *d de Cohen* (Cohen, 1988), mais il en existe d'autres comme Hattie (2009). L'ampleur de l'effet permet de comparer 2 moyennes : pré/post intervention, groupe contrôle/expérimental, etc. et d'estimer la taille de leur différence (McGough & Faraone, 2009). Selon Cohen, un *d* inférieur à 0.2 est interprété comme étant sans effet ; entre 0,2 et 0.5 comme un effet moyen/intermédiaire et plus grand que 0.8 comme un effet large (Cohen, 1988). Hattie (2009) interprètera un effet inférieur à 0,2 comme un effet développemental, un effet entre 0,2 et 0,4 comme un effet de l'enseignant et un effet supérieur à 0,4 comme un effet désiré (Baye & Dachet, 2022).

Il convient de considérer que les valeurs (absolues) inférieures à 0.2 (ou 0.15 selon l'auteur) reflètent une différence ténue, pouvant être due à des erreurs de mesure ou à d'autres phénomènes aléatoires. Il s'agira donc d'être prudent dans

l'interprétation des ampleurs de l'effet comprises entre -0,2 et +0,2. (Baye & Dacht, 2022, p. 13)

Les tailles d'effets obtenues peuvent être soit positives ou négatives en fonction du sens de l'hypothèse. Si l'intervention vise une meilleure performance d'une certaine variable dans le groupe expérimental, une taille positive est attendue.

Selon les données disponibles dans les études sélectionnées de cette revue, les tailles des effets ont été calculées à l'aide de trois méthodes différentes :

- La comparaison de deux groupes de taille identique :

Le groupe expérimental (bénéficiant d'une des quatre pratiques innovantes) et le groupe contrôle (bénéficiant d'un enseignement traditionnel) ont le même effectif. Le calcul s'effectue à l'aide des deux moyennes et de l'écart-type de chaque groupe. La taille de l'effet obtenue est alors appelée *d de Cohen*. Dans le cas de nos calculs, les écarts-types étaient parfois fort différents ; dans ce cas, il est recommandé d'utiliser l'écart-type du groupe contrôle, car celui-ci n'est normalement pas influencé par l'intervention.

- La comparaison de deux groupes de tailles différentes :

Lorsque les deux groupes présentent des échantillons de tailles différentes, le calcul est identique à la méthode précédente si ce n'est que l'écart-type regroupé est calculé en fonction des valeurs des tailles des échantillons. Cela permet de corriger un biais au niveau de l'écart-type combiné.

- La taille de l'effet pour des groupes de tailles différentes avec pré/post-test :

Quand une étude propose un design de pré-test/post-test « les différences dans les mesures pré-post dans les groupes doivent être pondérées par l'écart type regroupé de la pré-mesure car cela n'a pas été influencé par l'intervention » (psychometrica, s.d).

3.10. Interprétation de l'ampleur de l'effet moyen

Après avoir calculé les ampleurs de l'effet pour chaque variable, nous avons combiné les effets ayant un lien entre eux afin d'obtenir une ampleur de l'effet globale, aussi appelée : « ampleur de l'effet moyen ». Comme l'explique Borenstein et al. (cité par Baye et Dacht, 2022) si les ampleurs de l'effet sont cohérentes, nous pouvons mettre en évidence la puissance de l'effet global, si les ampleurs de l'effet ont une variance modérée, nous pouvons nous baser sur l'effet moyen en ayant conscience que les effets varient d'une étude à l'autre et si les ampleurs de l'effet varient de manière importante, nous devons plutôt analyser l'écart entre ces ampleurs afin de définir laquelle est vraie.

3.11. Critères de qualité des études incluses

Chaque étude a été analysée individuellement selon un tableau reprenant différents critères de qualité (Baye & Dachet, 2022) tels que : le design de recherche ; la taille des effectifs ; les outils de récolte de données sélectionnés par les chercheurs ; la durée de l'étude et l'équivalence au prétest. Cette analyse permet de détecter les risques de biais des études. Un tableau détaillé de chaque étude est disponible en annexe 1 et la cote finale obtenue est présenté dans le tableau présentant les différentes caractéristiques des études.

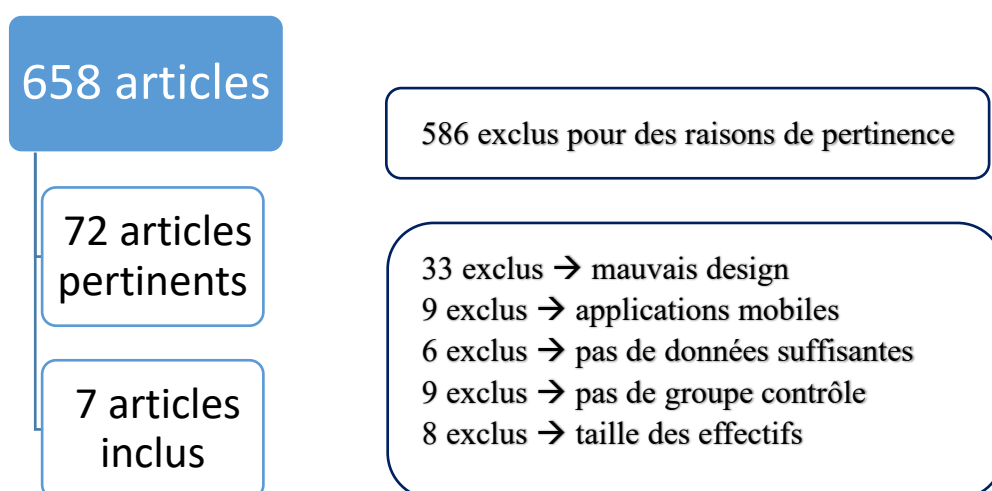
4. Présentation des résultats et discussions

Les différentes équations de recherches ont permis de recueillir 1989 références. Plus précisément, 1617 résultats ont été trouvés dans ERIC et 372 dans SCOPUS. Parmi celles-ci, 197 doublons ont été supprimés. La phase de sélection des articles a été menée sur 1792 titres et résumés. 1495 références comportaient un ou plusieurs critères d'exclusion, laissant 297 textes à lire de manière sélective ou en entier. Pour chaque étude sélectionnée, nous avons réalisé : Un flowchart PRISMA (Page et al., 2021) ; un tableau reprenant les caractéristiques de chaque étude ; une grille d'évaluation des études (Baye & Dachet, 2022) (annexe 1) ; un tableau reprenant les programmes utilisés dans chaque étude (annexe 2) et un tableau expliquant les différents outils de récolte de données (annexe 3). Pour un souci de clarté et de compréhension, la partie « résultats » d'une pratique innovante est directement suivie de la partie « discussion ».

4.1. Résultats et discussion : « École du dehors »

4.1.1. Flowchart (PRISMA, 2020)

Comme expliqué précédemment, dans cette revue, nous nous sommes intéressés essentiellement aux études montrant des effets sur les performances des élèves, peu importe la matière scolaire abordée. A contrario, les études centrées sur le bien-être, les émotions, la motivation ou les perceptions des élèves ou des enseignantes ont été exclues. Plusieurs études (N=12) combinaient l'apprentissage en dehors de l'école et la technologie mobile. Lorsque les effets de ces deux approches n'étaient pas dissociés, nous avons été contraints de les exclure également.



4.1.2. Caractéristiques des études incluses « école du dehors » (Baye & Dachet, 2022) :

N	Auteur(s)	Intitulé	Design	Effectif	Durée	Pays	Variable(s)	Effet(s)	Qualité
1.	Otte, C.R., Bølling, M., Stevenson, M.P., Ejbye- Ernst, N., Nielsen, G. & Bentsen, P. (2019)	Education outside the classroom increases children's reading performance: Results from a one-year quasi- experimental study.	Quasi- expérimental Pré-test Post test	Élèves de 8 à 13 ans GE : 376 GC : 153	5h/semain e pendant une année scolaire	Danemark	Performances en lecture	0,2	6/6
2.	Quibell, T. ; Charlton, J. ; Loi, J. (2017)	Wilderness Schooling: A controlled trial of the impact of an outdoor education programme on attainment outcomes in primary school pupils	Quasi- expérimental -Pré-test (T1) -Post test (T2) puis après 6 semaines (T3)	9 écoles Élèves de 8 à 11 ans GE: 223 GC: 217	1 jour/ semaine durant 6 semaines	Royaume- Uni	-les performances en lecture -les performances en rédaction -les performances en mathématiques	0,953 0,859 0,629	4/6
3.	Wells, N. M., Myers, B. M., Todd, L. E., Barale, K., Gaolach, B., Ferenz, G., Aitken, M., Henderson, C. R., Tse, C., Pattison, K. O., Taylor, C., Connerly, L.,	The Effects of School Gardens on Children's Science Knowledge: A randomized controlled trial of low- income elementary schools	Longitudinal: essai contrôlé randomisé	47 écoles 151 classes Élèves de 6 à 12 ans GC: 1439 GE: 1622	2 ans	États-Unis	Connaissances scientifiques	0,202	5/6

	Carson, J. B., Gensemer, A. Z., Franz, N. K., & Falk, E. (2015)								
4.	Ting, K.L.; Siew, N.M. (2014)	Effects of Outdoor School Ground Lessons on Students' Science Process Skills and Scientific Curiosity	Quasi-experimental	2 classes GE: 63 GC: 56	6 semaines	Malaisie	Compétences en matière de processus scientifiques	pré/ post GE 1,367 post-test GC/ GE 0,77	4/6
5.	Fancovicova, J. ; Prokop, P. (2011)	Plants have a chance: Outdoor educational programmes alter students' knowledge and attitudes towards plants	Quasi-experimental	10 à 11 ans GE: 17 GC: 17 élèves	3 x 120 minutes	Slovaquie	Connaissances des élèves à l'égard des plantes	Post test 3,24 Re test 3,14	2/6
6.	Chang, C-S. ; Chen, T-S. ; Hsu, W-H. (2011)	The Study on Integrating WebQuest with Mobile Learning for Environmental Education	Quasi-experimental	103 élèves 11 à 12 ans GC1: 34 GC2 : 34 GE : 34	1 mois	Taiwan	Performances d'apprentissage de l'éducation environnementale des élèves	-0,947 GC1/GE -0,338 GC2/GE	2/6
7.	Carrier, S. (2009)	Environmental Education in the Schoolyard: Learning Styles and Gender	Quasi-experimental Carrier, S. (2009)	Élèves de 11 à 12 ans GE: 27 GC: 25	14 semaines	États-Unis (Sud Est)	Connaissances environnementales pour les garçons	1,329	2/6

4.1.3. Description générale des études « école du dehors »

Sept études répondant à nos critères d'éligibilité ont été trouvées. Parmi ces études, celle de Wells et al. (2015, 3), utilise un design expérimental longitudinal avec essai contrôlé randomisé. Un processus de randomisation via un programme informatique est utilisé permettant d'assigner aléatoirement les participants aux différents groupes. Les six autres études (1, 2, 4, 5, 6, 7) suivent un design quasi-expérimentale. Ce type de recherche se différencie du design expérimental pur, notamment dans la manière, non aléatoire, de répartir les participants dans les groupes expérimentaux ou de contrôles. Le recours à un groupe témoin est primordial afin de garantir que des éléments externes à l'intervention n'affectent pas l'interprétation des données. Un pré-test est également nécessaire pour pouvoir comparer une variable dépendante avant et après l'intervention (Pinard et al., 2004).. Dans l'étude 5, les auteurs (Fancovicova, & Prokop, 2011) indiquent répartir les élèves au hasard mais renseignent l'étude comme étant quasi-expérimentale. Dans les autres études (1, 2, 4, 6 et 7) la répartition des élèves dans les groupes contrôles ou expérimentaux est effectuée par les enseignants en essayant de constituer des groupes les plus similaires possible. Ce processus est souvent utilisé en sciences de l'éducation lorsque la randomisation est impossible au niveau organisationnel. Dans l'étude 4, les classes entières sont réparties en groupe contrôle ou expérimental. Cependant, l'équivalence entre les groupes avant le début de l'intervention est évaluée dans toutes les études à l'aide de pré-tests. Dans l'étude 3, une différence au niveau de l'âge des participants est constatée mais reste minime ($GC=8,22$ $GE=8,34$). Dans l'étude 7, les groupes ne sont pas équivalents entre le pré-test et le post test, compromettant la validité de cette étude dans la mesure où il est difficile de définir si les résultats observés sont dus à la variable manipulée ou à ce changement dans les groupes. Il est important d'en tenir compte lors de l'interprétation des résultats.

Variables mesurées

Les variables d'intérêts des études sont diverses et variées. Les effets des programmes de l'école du dehors sont étudiés sur les performances : en langue maternelle ($k = 3$), en math ($k = 1$), en sciences ($k = 4$) et en connaissances environnementales ($k = 3$).

Les ampleurs de l'effet extraites ou calculées ont un effet faible à large (Cohen, 1988) en fonction des études et ont toutes un signe positif signifiant que les élèves bénéficiant d'une intervention « école du dehors » obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Cependant, ces résultats sont à interpréter avec beaucoup de

prudence en fonction de la qualité et les caractéristiques des études. Nous y reviendrons ultérieurement.

Programmes utilisés dans les études

Différents programmes sont étudiés dans les études proposées (un tableau détaillé comprenant les interventions dans les groupes expérimentaux et les groupes contrôles se trouve en annexe 2.A.). Ces derniers se rejoignent sur plusieurs points : Ils expérimentent tous un programme se déroulant en dehors des classes : parfois uniquement dans la cour de l'école (études 4, 6 et 7) ou dans un environnement extérieur pour les quatre autres études. Ces activités se déroulent au minimum une fois par semaine et sont majoritairement données par les enseignants participants à l'exception de l'étude 2 où le programme est dispensé par des professionnels appartenant à l'étude.

Cependant, ils se distinguent sur de nombreux points. En effet, aucune étude n'utilise le même programme ou ne se base sur les mêmes sources pour en créer un. Certains sont très précis (2, 3, 4, et 5) et d'autres sont plutôt des guides pour les enseignants désireux de s'engager dans une étude (1, 6 et 7). Les performances mesurées sont très variées au niveau des matières scolaires ou au niveau des outils. Enfin, trois études sur sept (1, 5 et 6) pratiquent une activité spécifique dans le groupe contrôle (annexe 3.A.). Cependant, cela reste minime comparativement au dispositif mis en place dans les groupes expérimentaux.

Déroulement des études

La durée de la plupart des études (2,4,5,6) ne dépasse pas 12 semaines et l'étude 7 se déroule pendant 14 semaines. Les effets du dispositif mis en place sont, par conséquent, analysés à court terme et ne peuvent pas garantir l'efficacité de l'intervention sur le long terme. Les deux seules études fiables au niveau de la durée sont celle de Otte et al. (2019, 1) analysant les effets sur les performances en lecture pendant un an ($d=0,2$) et celle de Wells et al. (2015, 3) qui évaluent les effets du dispositif sur les connaissances en sciences des élèves pendant deux ans ($d=0,202$). Il est primordial de mettre en exergue ces amplitudes de l'effet considérées comme sans effet ou très faible (Cohen, 1988), ne permettant pas d'établir un lien entre les performances des élèves et « l'école du dehors » sur le long terme.

Participants

La taille des effectifs est une qualité générale pouvant être attribué à pratiquement toutes les études présentées dans cette partie du rapport. En effet, le nombre de participants est toujours

supérieur à trente sujets. Plusieurs études (1,2,3,4 et 6) ont d'ailleurs un effectif élevé en testant plus d'une centaine de sujets, permettant d'obtenir des résultats plus représentatifs de la population (Abbadia, 2023). Les études 5 et 7 présentent des échantillons inférieurs à 30 élèves par groupe. Elles présentent des effets très forts (Cohen, 1988) entre l'école du dehors et les performances, ces résultats ne peuvent pas être généralisés à l'ensemble de la population. Pour rappel, l'étude 5 évalue les effets du dispositif sur les connaissances en sciences et l'étude 7 sur les connaissances environnementales.

Concernant l'âge des participants, il reste assez similaire entre les études, généralement entre 8 et 13 ans, permettant une comparaison plus aisée. Aucune étude n'a pour objet des enfants du premier degré de l'enseignement primaire, par conséquent, les effets des traitements étudiés ne peuvent pas s'appliquer à ce niveau scolaire.

Outils de récolte de données

Chaque étude utilise un outil de récolte de données spécifique (un tableau détaillé comprenant les différents outils de mesure se situe en annexe 3.A.). Tous les participants au sein d'une même étude sont évalués de la même manière, avec le même instrument sauf dans l'étude 6 où le pré-test et le post-test ont été conçus par les enseignants et ne sont pas identiques. Il est, dans ce cas, probable que les deux tests ne mesurent pas les performances en éducation environnementale de la même manière avant et après l'intervention. De plus, dans la majorité des études (2,4,5,7), les outils d'évaluation ont été créés par les chercheurs et sont donnés par les enseignants en classe et non pas un évaluateur indépendant. Ces différents points peuvent entraîner des biais dont nous discuterons lors de l'interprétation des résultats.

En outre, deux études (1 et 3) indiquent que les outils sont standardisés, ces études mesurent respectivement les performances en lectures et en sciences. Les autres ne précisent pas cette information. Pour rappel, ces études présentent des ampleurs de l'effet très faibles (Cohen, 1988).

Synthèse des résultats

Au regard de l'analyse et de l'évaluation de ces études, nous pouvons constater un manque de critères de qualité pour trois d'entre elles (5, 6 et 7). Un tableau détaillé se trouve en annexe 1.A. Entre une non-équivalence entre les groupes (7), un pré-test différent du post-test (6) et un manque d'informations (5), leurs résultats sont à interpréter avec beaucoup de prudence. Les deux meilleures études au niveau de la qualité sont celles de Otte, C.R et al. (2019, 1) et Wells et al. (2015, 3). Les auteurs utilisent des outils de collecte de données

standardisés, les effectifs sont élevés et les études durent minimum un an. Les performances en lecture et en sciences sont mesurées et indiquent des ampleurs de l'effet « sans effet » (Cohen, 1988). A contrario, deux autres études (2 et 4), comptabilisant plusieurs critères de qualité également, démontrent des effets moyens à forts de cette pratique sur les performances en français, en math et en sciences. Il est, par conséquent, difficile de voir un effet indiscutable de « l'école du dehors » sur les performances en analysant les meilleures études sélectionnées.

4.1.4. Discussion « école du dehors »

Les effets positifs de la nature sur la santé physique et mentale sont à l'heure actuelle bien connus. Nous savons aujourd'hui, à titre d'exemple, qu'être en contact avec des espaces verts réduit le taux de cortisol, l'hormone du stress et diminue la tension artérielle (Dulieux, 2022), mais nous en savons moins sur les processus d'apprentissage des enfants en contact avec la nature (Dadvand et al., 2015). D'après Kuo et al. (2019), les premières recherches sur le sujet manquaient cruellement d'objectivité et contenaient de nombreux biais. Certaines théories plausibles émergeaient toutefois comme celle de Kaplan et Kaplan (1989) qui attestaient que le contact avec la nature améliore les capacités cognitives par le biais « des ressources attentionnelles dirigées » (Clerquin & Sastourné-Arrey, 2019). D'après Kuo et al. (2019), à l'heure actuelle, les études scientifiques empiriques se multiplient dans des contextes variés en faisant preuve de plus en plus de rigueur. Par conséquent, comme expliqué précédemment, nous avons analysé les recherches existantes en centrant nos analyses uniquement sur les performances des élèves afin d'amener aux enseignants des moyens de lutter contre les faiblesses de notre système scolaire.

Pour rappel, nous avons sélectionné sept études répondant à nos critères d'éligibilité, nous avons, ensuite, extrait ou calculé les ampleurs de l'effet des programmes de « l'école du dehors » sur les performances : en langue maternelle ($k=3$), en math ($k=1$), en sciences ($k=4$) et en connaissances environnementales ($k=3$).

Postérieurement, nous avons calculé des ampleurs de l'effet moyen permettant de combiner les effets de plusieurs variables (Baye et Dachet, 2022). Plus l'ampleur de l'effet moyen est grande, plus l'effet étudié est considéré comme important (Roques, 2022).

Les ampleurs de l'effet moyen proposés ci-dessous présentent l'effet de « l'école du dehors » sur les variables mesurées ayant un lien entre elles :

Variables	Études utilisées	Ampleurs de l'effet moyen
Performances en langue maternelle (lecture et rédaction)	1 et 2	0,671
Performances en sciences	3,4 et 5	1,743
Performances en éducation environnementale	6 et 7	0,871
Performances toutes matières scolaires confondues	Toutes	1, 121

Les ampleurs de l'effet global calculées montrent des effet moyens à forts selon Cohen (1988) en fonction des performances mesurées et ont toutes un signe positif pouvant signifier que les élèves bénéficiant d'une intervention « école du dehors » obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Cependant, différents biais ont été relevés précédemment et nous allons tenter de mieux comprendre leur influence.

- **Performance en langue maternelle (lecture et rédaction)**

Ces ampleurs de l'effet moyen se basent sur deux études (Otte et al., 2019 ; Quibell et al., 2015) et sont estimées dans ce rapport comme étant de bonne qualité. Cependant, si l'étude de Otte et al. (2019) se déroule pendant une année entière, celle de Quibell et al. (2015) ne dure que 6 jours répartis sur 6 semaines. Cette durée, très courte, ne permet pas d'affirmer les bénéfices de l'école du dehors sur le long terme. De plus, cette étude ne donne pas beaucoup d'informations sur leurs outils de récolte de données contrairement à ceux utilisés dans l'étude de Otte et al. (2019) qui sont standardisés. Dès lors, il semble important de mettre en évidence que l'étude de Quibell et al. (2015) influence fortement la moyenne calculée. Si nous analysons les effets de l'école du dehors en retirant cette étude, l'ampleur de l'effet devient faible ($d=0,2$): « Il convient de considérer que les valeurs (absolues) inférieures à 0.20 (ou 0.15 selon l'auteur) reflètent une différence ténue, pouvant être due à des erreurs de mesure ou à d'autres phénomènes aléatoires » (Baye et Dachet, 2022, p. 13). Par conséquent, nous ne pouvons pas affirmer que « l'école du dehors » influence les performances en langue maternelle (lecture et rédaction) sur le long terme.

- **Performances en sciences**

L'ampleur de l'effet moyen se base sur trois études de qualité variables (3, 4 et 5). L'étude 3 (Wells et al, 2015) est considérée, d'après nos critères, comme une étude fiable ainsi que l'étude 4 (Ting et al., 2014) avec, cependant, un peu plus de nuance. L'étude 5 (Fancovicova et al., 2011) quant à elle, présente plusieurs facteurs discutables comme un effectif excessivement petit (17 participants par groupe). Il est probable que cet échantillon ne soit pas représentatif de

la population et, par conséquent, ne permet pas de généraliser les résultats. De plus, si « l'échantillon est trop petit, seuls les effets très marqués (ou assortis d'une très faible variabilité) pourront être décelés » (Degraeve, 2022, chap. 11). Si nous regardons les ampleurs de l'effet de cette étude séparément, nous pouvons constater des valeurs extrêmes ($d=3,24$ et $d=3,14$). L'analyse de sensibilité de l'ampleur de l'effet moyen de cette étude indique une surestimation des effets positifs. Si nous excluons cette étude, l'ampleur de l'effet moyen passe de 1,743 à 0,779, considéré comme étant un effet intermédiaire d'après Cohen (1988). Nous pouvons interpréter cette valeur comme une influence moyenne de l'école du dehors sur les performances en sciences.

- **Performances en éducation environnementale**

L'ampleur de l'effet global se base sur deux études (6 et 7) de qualité très discutable.

Dans l'étude 6 (Chang et al., 2011), le pré-test et le post-test ont été conçus par les enseignants et ne sont pas identiques. La probabilité que les deux tests ne mesurent pas les performances en éducation environnementale de la même manière avant et après l'intervention est à prendre en considération.

Dans l'étude 7 (Carrier, 2009), les groupes ne sont pas équivalents entre le pré-test et le post test, compromettant inévitablement la validité de cette étude dans la mesure où il est difficile de déterminer si les résultats observés sont dus à la variable manipulée ou à ce changement dans les groupes. De plus, cette étude présente des échantillons inférieurs à 30 élèves par groupe compromettant la généralisation des résultats comme expliqué précédemment.

Pour toutes ces raisons, il semble judicieux de ne pas tenir compte des résultats obtenus concernant les performances en éducation environnementale.

- **Performance en mathématiques**

Une seule étude évaluait les performances dans cette matière scolaire ne permettant pas de calculer une ampleur de l'effet global ou de tirer des conclusions généralisables. Par conséquent, d'autres recherches sont nécessaires afin d'établir un lien entre l'école du dehors et cette variable.

4.1.5. Conclusion des études pratiquant « l'école du dehors »

Après l'analyse des résultats des études en lien avec l'école du dehors, nous pouvons conclure que :

- les effets de cette pratique sur les performances en lecture et en rédaction ne sont pas suffisamment marqué pour affirmer une influence positive.

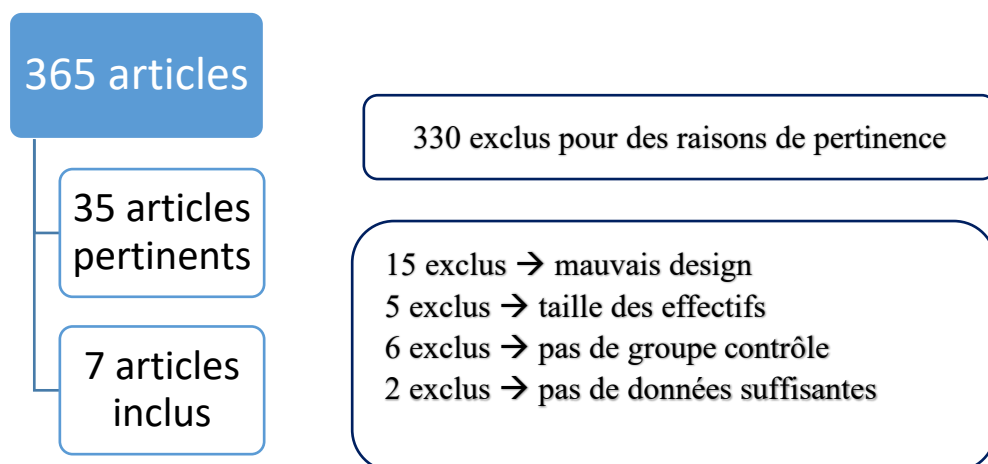
- L'école du dehors a un effet modéré sur les performance en sciences. Ceci peut éventuellement s'expliquer par les objectifs de cette matière scolaire centrés sur l'observation, l'expérimentation ou sur les végétaux et les animaux par exemple, sujets facilement exploitables en extérieur. Pour rappel, la population représentée dans nos analyses se situe entre 8 et 13 ans, par conséquent, nos constats ne s'appliquent pas aux enfants du premier cycle de l'enseignement primaire.
- Les effets sur l'éducation environnementale ne peut pas être établi suite à l'analyse des études sélectionnées dans ce rapport.

Bien que les études se multiplient ces dernières années, il semblerait qu'un manque de recherches rigoureuses figure dans la littérature scientifique concernant l'efficacité de cette pratique sur les performances. Seulement sept études accessibles répondaient à nos critères d'éligibilité et certaines d'entre elles présentaient une qualité discutable. En outre, plus de clarté et d'informations dans les études serait souhaitable afin de permettre aux différents lecteurs d'en évaluer les caractéristiques.

4.2. Résultats et discussion : Les cartes mentales

Lors de nos recherches, nous avons inclus les études utilisant les cartes mentales manuscrites ou via un logiciel, sans distinction. Nous avons exclu celles utilisant des schémas comme ceux pratiqués pour la résolution de problèmes en mathématiques, pour la théorie des schémas ou la résolution des problèmes de mots (WPS), ceux-ci ayant une procédure bien spécifique, différente de celle des cartes mentales.

4.2.1. Flowchart (PRISMA, 2020)



4.2.2 Caractéristiques des études incluses (Baye & Dachet, 2022)

N	Auteur(s)	Intitulé	Design	Effectif	Durée	Pays	Variable(s)	Effet(s)	Qualité
1.	Wan, H; Yu, S. (2023)	A recommendation system based on an adaptive learning cognitive map model and its effects.	Quasi-expérimental Post test uniquement	69 élèves de 11 ans GE : 34 GC : 35	1h ½ par semaine durant 4 semaines	Chine	Acquis d'apprentissage en ligne	0,97	2/6
2.	Fung, D.; Liang, T. (2023)	The Effectiveness of Collaborative Mind Mapping in Hong Kong Primary Science Classrooms	Quasi-expérimental Pré-test Post test	3 classes de 4 ^{ème} année GE : 75 GC : 75	12 semaines (3 x 90 minutes par semaine)	Chine (Hong Kong)	Compréhension des concepts scientifiques T1 : pas d'intervention T2 : collaboration T3 : collaboration + carte mentale	T1/T3 0,719 T2/T3 0,769	5/6
3.	Merchie, E.; Catrysse, L.; Van Keer, H. (2021)	Mind Maps as primers when Reading-for-learning in elementary grades? An eye tracking study	Quasi-expérimental	64 élèves de 11 à 12 ans GE (MMT): 23 GE (TMM) : 21 GC : 20	3 semaines	Belgique (Gand)	Performance en lecture compréhension et mémorisation	MMT = 0,478 TMM = 1,529	2/6
4.	Aliaser, A.M. (2017)	The effectiveness of electronic Mind Maps in developing academic achievement and the attitude towards learning English among primary school students	Quasi expérimental	30 élèves de 10 à 12 ans GE : 15 GC : 15	5 semaines	Arabie Saoudite	Réussite scolaire des filles en langue anglaise	1,6	3/6

5.	Lan, Y. J., Sung, Y. T., Cheng, C. C., & Chang, K. E. (2015)	Computer-Supported Cooperative Prewriting for Enhancing Young EFL Learners' Writing Performance	Quasi expérimental	54 élèves de 5 ^{ème} année GE : 27 GC : 27	6 semaines 40 minutes 2x semaine	Taiwan	-Performance en grammaire anglaise langue étrangère -Performance en rédaction anglaise langue étrangère	Pré/post 0,648 Pré/post 0,488	1/6
6.	Balim, A. G. (2013)	The effect of mind-mapping applications on upper primary students' success and inquiry-learning skills in science and environment Education	Quasi expérimental Pré-test Post test	64 élèves de 7 ^{ème} année GE : 32 GC : 32	4h par semaine durant 3 semaines	Turquie	Réussite scolaire en sciences	1,659	3/6

4.2.3. Description générale des études utilisant des cartes mentales

Six études répondant à nos critères d'éligibilité ont été trouvées. L'ensemble de ces études ont un design de recherche quasi-expérimentale. Bien que celui-ci ait tendance à surestimer l'ampleur de l'effet par rapport aux études randomisées (Cheun & Slavin, 2016), il reste le meilleur moyen d'observer des résultats de tendances générales lorsqu'une randomisation n'est pas possible comme souvent en sciences de l'éducation. Une des études (Merchie et al., 2021) a la particularité d'utiliser un suivi oculaire. Cette technique permet de comprendre comment les individus traitent et interprètent visuellement les informations qui leur sont présentées (Guillon, Hadjikhani & Rogé, 2014).

Même si les participants sont répartis dans les groupes contrôles ou expérimentaux de manière non-aléatoire, dans la majorité des études (2,3,4,5,6), des pré-tests permettent de vérifier si les groupes ont bien des caractéristiques similaires avant le début de l'intervention, ce qui est le cas systématiquement, sauf pour l'étude de Wan (2023, 1) qui n'utilise pas de pré-test. Sans randomisation ni pré-test, il est fort probable que les deux groupes ne soient pas équivalents. Comment savoir si les résultats obtenus ne reflètent pas cette différence à la place des effets de l'intervention ? Il est indispensable de tenir compte de ce manque de rigueur lors de l'interprétation des résultats.

Variables mesurées

Les variables d'intérêts des études utilisant les cartes mentales sont très variées. Les effets des programmes sont étudiés sur les performances : des acquis d'apprentissage en ligne ($k=1$), en sciences ($k=3$), en lecture compréhension et mémorisation ($k=2$), en anglais ($k=3$) en mathématiques ($k=1$)

Les amplitudes de l'effet extraites ou calculées (tableau 4.2.2) présentent des effets faibles à forts (Cohen, 1988) en fonction des études et ont toutes un signe positif pouvant signifier que les élèves bénéficiant d'une intervention utilisant une carte mentale obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Cependant, il est indispensable d'analyser la qualité et les caractéristiques des études afin d'interpréter les résultats.

Programmes utilisés dans les études

Différents programmes sont présentés dans les études proposées, un tableau détaillé comprenant les interventions dans les groupes expérimentaux et les groupes contrôles se trouve en annexe 2.B. Ces derniers se rejoignent sur plusieurs points : Ils expérimentent tous un

programme utilisant une carte mentale, la plupart se réalise via un logiciel sauf pour l'étude de Balim (2013, 6) où les cartes mentales sont réalisées manuscritement. Ces activités sont toutes données par les enseignants participants, excepté pour l'étude de Wan (2023,1) combinant des activités données par les enseignants et des formations données par les chercheurs. Et enfin, tous les programmes proposent une intervention minimum une fois par semaine.

Cependant, ils se distinguent sur de nombreux points. En effet, aucune étude n'utilise le même programme ou ne se base sur les mêmes sources pour en créer un. Les performances mesurées sont très variées au niveau des matières scolaires ou au niveau des outils. De plus, une étude (2) donne une formation sur les cartes mentales à tous les élèves, y incluant les élèves du groupe contrôle. Cette particularité peut avoir un impact sur les résultats car ceux-ci mesurent éventuellement plus l'efficacité de la pratique collaborative (également implémentée) jumelée à la carte mentale plutôt que l'efficacité de la carte mentale seule, ne permettant pas de généraliser les résultats. Concernant l'étude de Balim (2013, 6), le programme utilisé semble assez limité dans son implémentation. En effet, les activités proposées, concernant l'utilisation de cartes mentales, consistent à compléter une carte mentale inachevée. Soit ce dispositif est fort restreint, soit l'auteur ne donne pas suffisamment d'informations aux lecteurs pour leur permettre d'avoir une idée précise de l'intervention mise en place dans le groupe expérimental.

Déroulement des études

La durée des études n'est pas un critère de qualité attribuable à celles sélectionnées dans cette partie du rapport. En effet, aucune étude ne dépasse 12 semaines d'implémentation, signifiant que les effets de tous les dispositifs mis en place ne sont mesurés qu'à court terme. Il est donc important de prendre en compte cette limite lors de l'interprétation des résultats.

Participants

Concernant la taille des effectifs, même si dans trois études (1,2,6) le nombre de participants est supérieur à trente sujets, il ne l'est que très légèrement et aucune étude n'atteint plus de septante-cinq participants par groupe. Plus l'échantillon est petit, moins il permet d'obtenir des résultats représentatifs de la population. De plus, « un grand échantillon est également avantageux parce qu'il confère une puissance aux analyses statistiques, c'est-à-dire qu'il leur permet de révéler des différences ou des liens subtils » (Dion et al., 2022, p.105).

L'âge des participants reste assez similaire dans les différentes études, généralement entre 10 et 12 ans, facilitant une comparaison dans le dernier degré de l'enseignement primaire

mais ne permettant pas de l'utiliser avec tous les enfants de l'école primaire comme souhaité initialement dans cette revue.

Outils de récolte de données

Chaque étude utilise un outil de récolte de données spécifiques (un tableau détaillé comprenant les différents outils de mesure se situe en annexe 3.B.). Selon nos critères, seule une étude (2) est de qualité au niveau des outils de mesure renseignés comme étant indépendants, c'est-à-dire qu'ils n'ont pas été créés par le chercheur et cet instrument est standardisé. Dans les six autres études (1, 3, 4, 5, 6), soit l'outil a été créé par l'enseignant ou par le chercheur, soit l'information n'est pas renseignée. Selon de Boer et ses collègues (2014), un test auto-développé fournirait une taille de l'effet significativement plus élevée qu'un test indépendant de l'intervention de par la subjectivité qu'il peut contenir.

Dans l'étude 3, les récoltes des données se font via un suivi oculaire. Même si cette technique peut fournir des informations précieuses sur la manière dont les individus traitent visuellement l'information, la rapidité avec laquelle ils lisent et la manière dont leur attention visuelle est dirigée, ce processus peut présenter certains biais comme différents soucis techniques ou un inconfort pour les participants qui peut affecter la qualité des données recueillies (Guillon et al., 2014).

Synthèse des résultats

De manière générale, nous pouvons constater une certaine faiblesse au niveau de la qualité et de la fiabilité des études retenues. Seule une étude sur six (Fung, 2023, 2) répond suffisamment aux critères de qualité présentés en annexe 1.B. L'ampleur de l'effet de cette recherche montre que les élèves ayant assisté à une activité utilisant une carte mentale collaborative ont des performances en sciences supérieures aux deux autres groupes témoins. Cependant, pour rappel, cette étude recherchait à démontrer l'effet de l'utilisation de la carte mentale combiné au travail collaboratif et donnait une formation, à tous les étudiants, sur la construction de cartes mentales, en y incluant les élèves du groupe contrôle. Les deux ampleurs de l'effet relevées démontraient un effet modéré de l'utilisation de la carte mentale, cependant, dans cette conjoncture, il est difficile de savoir si les résultats obtenus sont imputables au dispositif mis en place ou à sa combinaison avec la pratique collaborative.

4.2.4. Discussion

Le concept de carte mentale est utilisé de plus en plus depuis le développement de la technologie (Cendros Araujo & Gadanidis, 2020) pour aider les enfants à visualiser et à organiser leurs connaissances (Fung et Liang, 2022). «La méthode heuristique (Mind Mapping) peut s'appliquer à tous les aspects de la vie où un apprentissage plus efficace et une réflexion plus claire peuvent améliorer les performances humaines.» (Buzan, 1995, cité par creativite.net, s.d.). D'après Faouzi et al. (2019), la carte mentale a démontré son efficacité dans plusieurs pays, notamment en Finlande et en France, en tant qu'outil pédagogique capable de favoriser l'apprentissage des élèves. Elle contribuerait, de manière probante, à l'amélioration de l'intelligence, de la créativité, de la communication, de la concentration et de la mémoire. Cependant, cette approche est systémique et il n'est pas si simple d'utilisation dans les premiers moments. De plus, certaines personnes travaillent plus efficacement avec des schémas linéaires et n'ont pas d'avantages à utiliser cette méthode (Rajteric, 2024).

A notre niveau, comment pouvons-nous interpréter les résultats de cette revue ? Pour rappel, nous avons sélectionnés six études répondant à nos critères d'éligibilité, nous avons, ensuite, extrait ou calculé les ampleurs de l'effet des programmes de « cartes mentales » sur les performances : des acquis d'apprentissage en ligne ($k=1$), en sciences ($k=3$), en lecture compréhension et mémorisation ($k=2$), en anglais ($k=3$) en mathématiques ($k=1$)

Ensuite, nous avons calculé des ampleurs de l'effet moyen permettant de combiner les effets de plusieurs variables (Baye et Dachet, 2022). Plus l'ampleur de l'effet moyen est grande, plus l'effet étudié est considéré comme important.

Les ampleurs de l'effet moyen proposés ci-dessous présentent l'effet de « l'utilisation de cartes mentales » sur les variables mesurées ayant un lien entre elles :

Variables	Études utilisées	Ampleurs de l'effet moyen
Performances en sciences	2 et 6	1,049
Performances en anglais 2 ^{ème} langue	4 et 5	0,912
Performances en lecture compréhension et mémorisation	3	1,003
Performances toutes matières scolaires confondues	Toutes	0,984

Les ampleurs de l'effet global calculées sont toutes considérées comme fortes (Cohen, 1988) et ont toutes un signe positif pouvant signifier que les élèves bénéficiant d'une

intervention utilisant une carte mentale obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Cependant, comme explicité précédemment, les études sélectionnées présentent de nombreux biais que nous analysons ci-dessous.

- **Performances en sciences**

L'ampleur de l'effet moyen des performances en science se base sur deux études et trois variables (Fung, 2023, 2 ; Balim, 2013, 6) et est considérée comme forte selon Cohen (1988). L'étude de Fung (2023) est une étude de bonne qualité d'après nos critères (Annexe 1.B.) surtout au niveau des outils de récolte de données provenant de TIMSS (Trends in Mathematics and Science Study), une étude internationale reconnue consacrée aux mathématiques et aux sciences visant à comparer les systèmes éducatifs de différents pays et utilisant des instruments de mesure standardisés. Ce dispositif utilise des mesures précises et reproductibles selon des normes et des procédures bien définies et est utilisé pour garantir la fiabilité des mesures et faciliter la comparaison des résultats (Ordre des Psychoéducateurs et Psychoéducatrices du Québec, 2022). Cependant, dans cette étude, une formation sur les cartes mentales est donnée à tous les élèves, y compris ceux du groupe contrôle. Cette particularité peut avoir un impact sur les résultats et entraîner un biais d'intervention simultanée (Bibliothèque ULaval, s. d.). De plus, cette étude recherche les effets de l'utilisation d'une carte mentale combiné au travail collaboratif. De ce fait, il semble difficile de comparer ces effets à d'autres dispositifs où la carte mentale est utilisée seule.

Concernant la deuxième étude utilisée (Balim, 2013) pour calculer l'ampleur de l'effet moyen en sciences, nous pouvons constater qu'elle ne remplit que trois critères de qualité sur six (Annexe 1.B.). La durée de l'étude ne dépasse pas 12 semaines, ne permettant pas de voir les effets du dispositif sur le long terme et les outils de récolte de données ne sont ni indépendants ni standardisés et peuvent, par conséquent, entraîner des biais dans les résultats comme le biais de confirmation. Effectivement, les chercheurs peuvent influencer inconsciemment les données collectées ou interpréter les résultats de manière subjective. De plus, l'absence de standardisation peut rendre inexact la comparaison des résultats entre différentes études même si ce n'est pas systématique (Goëta, & Ho-Pun-Cheung, 2022). Pour ces différentes raisons, nous ne pouvons tirer de conclusion fiable concernant les performances en sciences.

- **Performances en anglais 2^{ème} langue**

Cette ampleur de l'effet moyen considéré comme étant un effet large (Cohen, 1988) se base sur deux études de faible qualité (4 et 5) notamment au niveau des effectifs ne dépassant pas 30 sujets par groupe. Comme expliqué plus tôt, si la taille de l'échantillon est trop petite,

cela peut compromettre la validité des résultats et influencer les conclusions de l'étude car seuls les effets variants faiblement ou fortement pourront être perçus (Degraeve, 2022). Nous pouvons constater également que ces études sont de courtes durées, 5 et 6 semaines d'implémentation, révélant, à nouveau, un manque de preuves de l'efficacité à long terme. En supplément, les instruments de récolte de données ne sont pas standardisés impliquant les travers discutés précédemment.

Au vu de ces différentes faiblesses, nous pensons qu'une analyse de sensibilité de l'ampleur de l'effet moyen n'est pas intéressante car aucune étude mesurant les performances en anglais 2^{ème} langue ne semble fiable.

- **Performances en lecture compréhension et mémorisation**

L'ampleurs de l'effet moyen de ces deux variables se basent sur la même étude (Merchie et al., 2021, 3) et est estimée, dans ce rapport, comme étant d'une qualité discutable également (Annexe 1B.). Dans cette étude, les faiblesses sont multiples : que ce soit au niveau de la taille des effectifs trop petite, l'outil de récolte de données crée par le chercheurs ou la durée de l'étude de seulement trois semaines. Ce dernier point peut être compréhensible car cette étude récolte les données via un suivi oculaire. Cette technique peut fournir des informations intéressantes sur la manière dont les individus traitent visuellement l'information, la rapidité avec laquelle ils lisent et la manière dont leur attention visuelle est dirigée (Guillon et al., 2014), mais, pour rappel, ce processus peut être inconfortable pour les participants et affaiblir la qualité des données recueillies. De plus, se fier aux résultats d'une seule et même étude ne permet pas d'obtenir des preuves solides. Pour ces diverses raisons, nous ne pouvons pas tenir compte des résultats obtenus en lien avec les performances en lecture.

- **Performance en mathématiques**

Comme pour les recherches sur les études impliquant l'école du dehors, une seule étude évaluait les performances dans cette matière scolaire. Ce manque de résultats ne permet pas de calculer une ampleur de l'effet moyen ou de tirer des conclusions généralisables. Par conséquent, d'autres recherches sont nécessaires afin d'établir un lien entre l'utilisation de cartes mentales et cette variable.

4.2.5. Conclusion des études impliquant une carte mentale

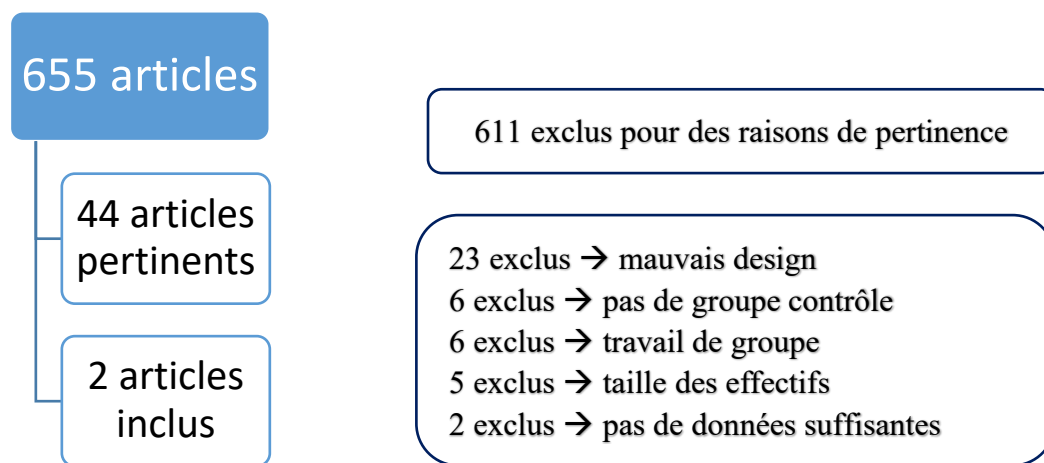
Après l'analyse des résultats de ces études, nous ne pouvons, malheureusement, pas tirer de conclusions généralisables et renseigner de manière fiable les enseignants désireux d'utiliser cette pratique. Trop de biais ont été relevés dans les études sélectionnées pour pouvoir amener des données probantes. Toutefois, de nombreuses études figurent dans la littérature scientifique

sur le sujet mais aucune ne répond à nos critères d'éligibilité ou ne communique suffisamment d'informations (Meter & Yokoi, 2007 ; Kalyuga et al., 2001 ; Farrand et al., 2002 ; D'Antoni et Pinto-Zipp, 2006 ; Karpicke et Blunt, 2011 ; Buzan, 2000). Il est nécessaire de souligner le manque de recherches scientifiques de qualité mesurant l'effet de cette pratique, sur les performances des élèves de l'enseignement primaire ordinaire.

4.3. Résultats et discussion : La classe flexible

Nous nous sommes centrés sur les effets que peut avoir l'environnement flexible de l'espace classe sur les performances des élèves toutes matières confondues. Nous n'avons pas retenu les études traitant de l'activité physique en classe ni l'utilisation des pauses actives, car, bien que ces activités soient liées à la mobilité que propose les classes flexibles, elles n'en font pas systématiquement partie.

4.3.1. Flowchart (PRISMA, 2020)



4.3.2. Caractéristiques des études incluses (Baye & Dachet, 2022)

N	Auteurs	Intitulé	Design	Effectif	Durée	Pays	Variable(s)	Effet(s)	Qualité
1.	Hardjansyah, F.; Misbahudhola m Abuyamin, M. (2022)	Enhancing Students' Learning Motivation through Changing Seats in Primary School	Quasi expérimental (pas de pré-test)	74 élèves de 4 ^{ème} année GE : 37 GC : 37	Information non communiquée	Indonésie	Résultats d'apprentissage	0,698	2/6
2.	Stapp, A. (2018)	Alternative seating and students' perceptions implications for the learning environment	Quasi expérimental	25 élèves de 10 à 11 ans Mesures répétées avec un groupe (pas de GC)	9 x 30 minutes durant 6 semaines	États-Unis	Temps à la tâche Temps hors tâche	1,16 -0,85	2/6

4.3.3. Description générale des études concernant la classe flexible

Seulement deux études répondant à nos critères d'éligibilité ont été trouvées. Chacune présente un design quasi-expérimentales. Cependant, la première étude (Hardjansyah, 2021) ne réalise pas de pré-test, dès lors, l'équivalence des groupes ne peut pas être assurée et il n'est pas possible d'évaluer réellement si l'exposition au dispositif a un effet. Un tableau détaillé de la qualité de chaque étude est disponible en annexe 1.C.

Un manque flagrant de recherches sur le sujet est à mettre en exergue concernant cette pratique innovante. Le dispositif étant relativement récent, cela explique certainement l'insuffisance des données.

Programmes proposés

Les programmes proposés dans les classes flexibles sélectionnées ne sont pas très développés ou les auteurs ne mentionnent pas suffisamment d'informations (un tableau détaillé est disponible en annexe 2.C.). En effet, les deux études proposent soit une disposition des sièges différentes de l'enseignement traditionnel (1) soit l'utilisation de coussins en forme de rond (2). La classe flexible est pourtant un ensemble « de principes fondamentaux : le choix, le mouvement, le confort, la vie en communauté et la coopération » (Allam & Chartier, 2023, p.132).

Variables mesurées

Les variables d'intérêts des études sont très limitées. Les effets des programmes d'assises flexibles sont étudiés sur les performances toutes matières confondues ($k = 1$), sur le temps à la tâche ($k=1$) et le temps hors tâche ($k = 1$).

Les ampleurs de l'effet extraites ou calculées ont un effet intermédiaire (Cohen, 1988) pour l'étude de Hardjansyah (2021) et fort pour l'étude de Stapp (2018) Les variables positives concernent les résultats d'apprentissage et le temps à la tâche ont un signe positif pouvant signifier que les élèves bénéficiant d'une classe flexible obtiennent de meilleures performances que les élèves étant dans une classe traditionnelle. L'ampleur de l'effet lié au temps hors tâche est négative, indiquant que les élèves se trouvant sur des sièges alternatifs passe moins de temps hors tâche que les élèves étant sur des sièges classiques. Cependant, comme cité précédemment, le manque d'études trouvées et leur qualité présentent beaucoup de biais.

Déroulement des études

Nous pouvons constater une durée trop courte d'implémentation du dispositif pour une étude (2) et un manque d'informations pour l'autre (1). Ceci ne permet dès lors pas de mesurer correctement tout l'impact que les classes flexibles pourraient avoir sur le rendement scolaire des élèves à long terme.

Participants

Un biais que l'on peut attribuer à ces études est une taille trop restreinte des effectifs. Comme Slavin et Smith (2009) le suggèrent, les études réalisées avec de grands échantillons apporteraient une estimation plus fiable et reproductible que des études à effectif réduit. Malheureusement, dans le cadre de cette recherche, trop peu d'études correspondant aux critères de sélection ont été trouvées.

Outils de récolte de données

Les deux études possèdent également une qualité discutable au niveau des outils de mesure utilisés (un tableau détaillé comprenant les différents outils de mesure se situe en annexe 3.C.). Peu d'informations sont transmises aux lecteurs et ne permettent pas de se faire une idée claire des méthodes utilisées. Un des outils a été créé par les chercheurs (Hardjansyah, 2022) et fournit probablement une taille de l'effet significativement plus élevée qu'un test indépendant de l'intervention. De plus, aucun outil de récolte de données n'est standardisé.

Synthèse des résultats

Au regard de l'évaluation de ces études, il est évident que trop de biais impactent la qualité et la fiabilité de ces recherches, ne permettant pas de tirer des conclusions sérieuses. De plus, le dispositif en soi peut amener certains débats parfois mouvementés comme expliqué ci-dessous.

4.3.4. Discussion

Au vu du peu de données récoltées pour cette pratique innovante et la faible qualité des deux études sélectionnées, il est difficile d'interpréter l'ampleur de l'effet moyen. En se référant à celle obtenue (effet fort, Cohen, 1988), nous pourrions considérer que le groupe utilisant des sièges alternatifs performe plus que le groupe utilisant des sièges traditionnels. De plus, d'après Stapp (2018), le temps « hors tâche » lors de l'utilisation de ces assises diminuerait fortement. Toutefois, les deux études sélectionnées présentent trop de lacunes non négligeables comme l'absence de pré-test pour la première ou l'absence de groupe contrôle pour la deuxième. Par conséquent, la discussion des résultats ne peut en aucun cas être généralisable à l'ensemble des dispositifs.

Et outre, la discutabilité des études ou le manque de données récoltées, le dispositif en soi peut également amener son lot de controverses. En effet, que signifie exactement un environnement de classe flexible ? Utiliser des assises un peu différentes suffit-il ? Selon Merrill (2018), proposer aux élèves une classe flexible signifie offrir un environnement dynamique et diversifié. La classe flexible se différencie de la classe traditionnelle par

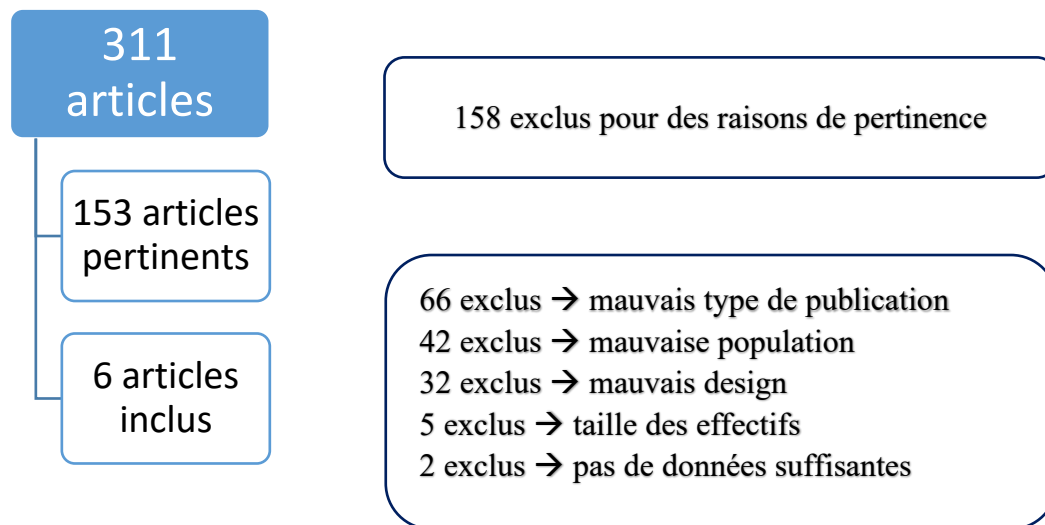
l'implication active des élèves dans leurs apprentissages et par la responsabilité de choisir une place de travail qui leur convient (Laquerre, 2018). Le but est d'offrir une variété de places pour permettre à l'élève de travailler dans la posture qu'il désire, mais aussi sur la surface de travail qu'il préfère (Delzer, cité par Laquerre, 2018). Les tables et les chaises peuvent être adaptées à différentes hauteurs, des coussins, des chaises à roulettes ou des ballons de stabilité peuvent également être proposés (Havig, cité par Leroux, 2021). Ces précisions permettent de se faire une idée plus précise de ce qu'est l'environnement de la classe flexible. Cependant, dans les études sélectionnées pour ce travail, la variable « enseignement » n'a pas été prise en compte. Prenons l'exemple d'un enfant qui dispose de tous les aménagements nécessaires à un environnement flexible, si cet enfant reçoit un enseignement magistral ou, à l'inverse, un enseignement plaçant l'enfant au centre des apprentissages, l'effet de l'environnement flexible aura-t-il le même impact ?

Dans un autre ordre d'idées, sachant que la classe flexible est un environnement d'apprentissage nouveau pour les élèves, il est possible qu'elle les déstabilise. Que dire alors des études implémentées pendant moins d'une année scolaire ? Pouvons-nous réellement y voir des effets tolérés ?

4.4. Résultats et discussion : Éveil aux langues

Lors de l'exploration des études existantes sur le sujet, nous avons mis l'accent sur l'éveil précoce aux langues étrangères. Afin d'accéder à un maximum d'études, nous avons élargi nos recherches de manière à inclure les élèves des écoles maternelles. Nous avons également accepté des études mesurant les effets d'un dispositif déjà en place. Par conséquent, certains programmes n'ont pas été implémentés par les chercheurs.

4.4.1. Flowchart (PRISMA, 2020)



4.4.2. Caractéristiques des études incluses (Baye & Dachet, 2022) :

N	Auteurs	Intitulé	Design	Effectif	Durée	Pays	Variable(s)	Effet(s)	Qualité
1.	Agegnehu, A.Z. ; Bachore, M.M.; Ayele, Z.A. (2023)	Effects of an explicit rime-based phonics method on the phonological awareness outcomes of Ethiopian EFL children	Quasi expérimental Pré-test Post test	GE: 35 GC: 35	2h/ semaine pendant 12 semaines	Éthiopie	PAT (test de conscience phonologique L2 mais aussi L1)	2,18	4/6
2.	Jaekel, N., Schurig, M., Van Ackern, I., & Ritter, M. (2022)	The impact of early foreign language learning on language proficiency development from middle to high school	Étude longitudinale Pré-test (année5) Post test (Année 7) Retest (année 9)	GE : 714 (débutants précoces : L2 en 1 ^{ère} primaire) GC : 805 (débutants tardifs: L2 en 3 ^{ème} primaire)	6 ans	Allemagne	- <i>Compréhension orale</i> -Compréhension écrite :	<i>Année 5 : 0,294</i> <i>Année 7 : 0,080</i> <i>Année 9 : 1,259</i> Année 5 : 0,314 Année 7 : 0,387 Année 9 : 0,516	6/6
3.	Wawire, B. A., & Kim, Y. G. (2018)	Cross-Language Transfer of Phonological Awareness and Letter Knowledge: Causal Evidence and Nature of Transfer	Expérimental Pré-test Post test	GE: 165 GC: 157 1ère primaire	8 semaines	Kenya	Transfert entre le kiswahili et l'anglais dans : - la connaissance du sons des lettres - la détection des phonèmes	0,62 0,95	5/6

							- la tâche de mélange conscience phono -la connaissance du nom des lettres -la tâche de lecture orale	0,37 0,27 0,00 ≤ d ≤ 0,08	
4.	Buser, M. (2020)	Oral proficiency development of K-4 learners of the Swiss two-way immersion program FiBi (<i>Filière Bilingue</i>) in a highly multicultural context	Étude longitudinale t1= début de la 1 ^{ère} primaire t2 : fin de 1 ^{ère} t3 : fin de 2 ^{ème} t4 : fin de 3 ^{ème} t5 : fin de 4 ^{ème}	GE : 868 (bilingues) GC : 305 (monolingues)	4 ans	Suisse	Les évaluations des compétences d'expression orale en allemand et en français	T1 : 9,95 T2 : 10,09 T3 : 7,96 T4 : 1,91 T5 : 2,87	3/6
5.	Maluch, J. T., Neumann, M., & Kempert, S. (2016)	Bilingualism as a resource for foreign language learning of language minority students? Empirical evidence from a longitudinal study during primary and secondary school in Germany	Étude longitudinale Pré-test Post test	-Monolingue (n = 820) -Allemand à dominante mixte (n = 119) - Non allemand à dominante mixte (n = 63) -Non Allemand uniquement (n = 30)	2 ans	Allemagne	Réussite en anglais (L3) entre la 6 ^{ème} année et la 8 ^{ème} année -Monolingue (n = 820) -Allemand à dominante mixte (n = 119) - Non allemand à dominante mixte (n = 63) -Non Allemand uniquement (n = 30)	0,56 0,55 0,42 0,27	4/6

6.	Lourenço, M., & Andrade, A. I. (2013)	Promoting Phonological Awareness in Pre-Primary Education: Possibilities of the "Awakening to Languages" Approach	Quasi experimental Pré-test Post test	Enfants de 3 à 6 ans GE : 21 GC : 21	1 an	Portugal	Développement de la conscience phonologique	0,934	4/6
----	---------------------------------------	---	---	--	------	----------	--	--------------	-----

4.4.3. Description générale des études : éveil aux langues

Six études répondant à nos critères d'éligibilité ont été trouvées. Ces études sont toutes des recherches quasi-expérimentales, cependant, plusieurs sont également longitudinales (2, 4, 5). Les effets d'une pratique comme l'éveil aux langues se mesurent plus efficacement sur le long terme : « Le choix du type d'études à inclure se fait en tenant compte du type d'études disponible et de la pertinence de celui-ci par rapport aux questions de recherche » (Sacré et al., 2021, p.7). Compte tenu des critères de qualité choisis dans ce travail (Baye & Dachet, 2022), nous pouvons considérer que les études présentées semblent fiables (Un tableau détaillé se trouve en annexe 1.D.). Dans l'étude 1 et 3, bien que les écoles aient été choisies par les chercheurs, les enfants, eux, ont été sélectionnés à l'aide de techniques d'échantillonnage probabiliste, cela signifie qu'ils ont été choisis au hasard (Statistique Canada, 2021). Cette procédure d'assignation aléatoire stratifiée permet d'éviter des biais d'échantillonnage. Dans les autres études (2,4,5 et 6), l'équivalence entre les groupes contrôles et expérimentaux est évaluée à l'aide de pré-tests pouvant vérifier si les groupes ont bien des caractéristiques semblables avant le début de l'intervention. Cependant, les études 4 et 6 présentent des particularités au niveau des groupes pouvant influencer les résultats. Le programme d'immersion mis en place dans l'étude de Buser & Melfi (2020, 4) intègre des apprenants francophones, germanophones et allophones. Hors, les apprenants allophones sont plus nombreux dans le groupe témoin, il est, par conséquent, possible que lors des premières années, les élèves allophones du groupe témoin ne connaissent pas leur langue de scolarisation, pouvant, de la sorte, orienter les résultats obtenus. Concernant l'étude 6 (Lourenço & Andrade, 2013), l'origine linguistique des élèves a peut-être pesé sur les conclusions car l'équivalence entre le groupe expérimental et témoin n'était pas respectée au niveau du taux d'immigrés (30 % dans le groupe expérimental et seulement 4,8% dans le groupe témoin).

Programmes proposés

Les programmes utilisés dans ces études sont divers et variés. Un tableau détaillé se trouve en annexe 2.D. Nous pouvons retrouver des formations sur la phonologie d'une langue étrangère (1, 3, 6) ou des cours précoces de langues étrangères (2). Néanmoins, toutes les études n'implémentent pas un programme spécifique dans cette partie du rapport. L'étude de Buser & Melfi (2020, 4) analyse un dispositif déjà en place : le programme d'immersion bidirectionnelle présent dans les écoles publiques suisses (BiFi) tandis que l'étude de Maluch et al. (2016, 5) mesure les effets du bilinguisme ou du monolinguisme à la maison.

Dans toutes les études, les élèves des groupes contrôles n'ont suivi aucun programme spécifique permettant de mieux mesurer les effets du traitement sur les groupes expérimentaux.

Variables mesurées

Les effets des programmes d'éveil aux langues sont étudiés sur les performances : en phonologique ($k = 6$), en compréhension orale ($k = 3$), en compréhension écrite ($k = 3$), en lecture ($k = 1$), les compétences linguistiques en allemand et en français ($k = 5$).

Les ampleurs de l'effet extraites ou calculées sont très variées (faible à fort selon Cohen, 1988) en fonction des études et ont toutes un signe positif signifiant que les élèves bénéficiant d'une intervention d'éveil aux langues obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Cependant, ces résultats sont à interpréter avec beaucoup de prudence en fonction des biais relevés dans certaines études. Nous en discuterons ultérieurement lors de l'interprétation des résultats.

Déroulement des études

La durée de la plupart des études (2,4,5,6,7) dépasse largement 12 semaines dû à l'inclusion d'études longitudinales. Les effets du dispositif mis en place sont analysés à long terme, ce qui est très intéressant, particulièrement pour cette pratique innovante. Les activités d'éveil aux langues reçues en maternelle ou en primaire permettent-elles d'avoir un impact positif sur l'apprentissage des langues dans le futur ?

Participants

La taille des effectifs est, également, une qualité générale que l'on peut attribuer pratiquement à toutes les études, le nombre de participants étant supérieur à trente sujets. Plusieurs d'entre elles (2,3,4,5) ont d'ailleurs un effectif très élevé comprenant plusieurs centaines de sujets. Par contre, l'étude 6 présente des échantillons inférieurs à 30 élèves par groupe, ne permettant pas d'obtenir des résultats suffisamment représentatifs de la population ou d'obtenir des valeurs incorrectes.

Concernant l'âge des participants, pratiquement toutes les classes de l'école primaire sont représentées. Une des études (6) commence le pré test en maternelle et une autre (5) réalise le post test en secondaire.

Outils de récolte de données

Une autre qualité accordée à presque toutes les études (1, 2, 3, 5, 6) est l'utilisation d'un outil de collecte de données indépendant. Cette étape est extrêmement importante dans les activités de recherche et permettent aux chercheurs de valider des résultats fiables. Néanmoins, la standardisation des données n'est réalisée que dans trois études (2,3,5). L'étude 4, quant à elle, utilise un outil de mesure créé par le chercheur, qui, de surcroît, évalue de l'expression orale. Lors de l'évaluation d'expression orale, plusieurs biais peuvent agir, comme le biais de confirmation, qui consiste à chercher des informations qui confirment des croyances existantes

plutôt que de voir celles qui les contredisent (Biais cognitif, 2024), le biais de stéréotype qui consiste à attribuer des caractéristiques spécifiques à un individu ou à un groupe d'individus sur la base de généralisations et de croyances souvent négatives (Fiske & Taylor, 2011) ou encore le biais de familiarité qui signifie que les chercheurs montrent une préférence pour les personnes ou les situations qui leur sont familières par rapport à celles qui sont moins familières ou inconnues (Tisseron, 2017). Ces biais peuvent amener à prendre des décisions précipitées, à mal interpréter les situations ou à avoir des réactions émotionnelles excessives.

Synthèse des résultats

Suite à l'analyse et l'évaluation de ces études, nous pouvons constater qu'elles ont été, majoritairement, menées de manière rigoureuse en limitant les biais. Cependant l'étude 4 (Buser, 2020) présente plusieurs faiblesses : un groupe contrôle et expérimental non équivalents, un outil de mesure non standardisé et non indépendant des chercheurs pour mesurer de l'expression orale. A contrario, une étude longitudinale (Jaekel et al., 2022, 2) se distingue par son excellente qualité. Celle-ci compare la compréhension orale et écrite, en anglais, d'enfants ayant eu des cours d'éveil aux langues précocement (première primaire) à des enfants ayant eu ces cours plus tard (3^{ème} primaire). Nous tiendrons compte de ces différents points lors de l'interprétation des résultats ci-dessous.

4.4.4. Discussion

Selon Porter (2019), nous pensons que les enfants ont de meilleurs résultats en matière de compétences linguistiques lorsque la première langue étrangère est enseignée dès le plus jeune âge. En effet, il est largement répandu que les jeunes enfants sont plus doués pour l'apprentissage des langues et qu'un commencement précoce permet un apprentissage plus long.

« Malgré les croyances dominantes, les jeunes apprenants en langues ont tendance à ne pas montrer d'avantages évidents en termes de taux d'apprentissage par rapport aux débutants tardifs, sans doute en raison de la cognition supérieure des apprenants plus âgés et d'un apport insuffisant pour déclencher les mécanismes d'apprentissage implicites préférés par les jeunes apprenants ». (Porter, 2019, introduction).

Fausse idées, croyances fondées, nous allons tenter d'y voir plus clair en interprétant les différents résultats récoltés.

Pour rappel, nous avons sélectionnés six études répondant à nos critères d'éligibilité. Nous avons, ensuite, extrait ou calculé les ampleurs de l'effet des programmes de l'éveil aux langues sur les performances : en conscience phonologique ($k = 6$), en compréhension orale

(k=3) en compréhension écrite (k=3), en lecture (k=1), des compétences linguistiques en allemand et en français (k = 5).

Postérieurement, nous avons calculé des ampleurs de l'effet moyen permettant de combiner les effets de plusieurs variables (Baye et Dachet, 2022). Plus l'ampleur de l'effet moyen est grande, plus l'effet étudié est considéré comme important (Roques, 2022).

Bien que les études incluses soient difficilement comparables entre elles, de par la diversité des dispositifs mis en place et de leur design, nous avons calculé des ampleurs de l'effet moyen sur des variables ayant un lien. Nous avons sélectionné celles relatives à :

Variables	Études utilisées	Ampleurs de l'effet moyen
Performances en conscience phonologique	1, 3 et 6	0,887
Performances en compréhension orale	2	0,544
Performances en compréhension écrite	2	0,405

Excepté pour la conscience phonologique, les études ne mesuraient pas les performances pour des matières scolaires similaires. Cependant, les ampleurs de l'effet moyen calculées sont considérées comme intermédiaire à large (Cohen, 1988) et ont toutes un signe positif pouvant signifier que les élèves bénéficiant d'une intervention utilisant l'éveil aux langues obtiennent de meilleures performances que les élèves n'ayant suivi aucune activité spécifique. Néanmoins, ces calculs incluent des études de pays différents et impliquent des langues dissemblables dans leur mode de fonctionnement, de plus, chaque dispositif et évaluation mis en place sont spécifiques.

- **Performances en phonologie**

En effet, prenons l'exemple de l'ampleur de l'effet moyen pour la variable « conscience phonologique » ayant été calculée à partir de trois études, il est notable d'insister sur le fait que, bien que les effets additionnés soient tous en lien avec la conscience phonologique, ils sont très différents dans leur implémentation. Le programme utilisé dans la première étude (Agegnehu, 2023, 1) concerne un enseignement explicite basé sur les rimes, un programme très spécifique peu utilisé dans des cours de langues étrangères classiques. La deuxième étude utilisée (Wawire & Kim, 2018, 3) se déroule au Kenya, un pays réputé pour sa société multilingue où l'on parle pas moins de soixante-sept langues ! Les enfants abordent-ils les langues étrangères de la même façon dans les pays monolingues ? Sont-ils aussi motivés pour évoluer dans ces cours ?

- **Performances en compréhension orale et écrite**

Les deux autres ampleurs de l'effet moyen ont été calculées sur base de la même étude (2). Pour rappel, cette étude montre une excellente qualité selon nos critères et mesure la compréhension orale et écrite, en anglais, d'enfants ayant eu des cours d'éveil aux langues précocement (première primaire) à des enfants ayant eu ces cours plus tard (3^{ème} primaire). Les deux ampleurs de l'effet moyen sont interprétées comme étant intermédiaires (Cohen, 1988), cela peut vouloir dire que le dispositif mis en place impacte moyennement les performances en compréhension des élèves. Si nous analysons de plus près les données, nous pouvons remarquer que cette étude montre des ampleurs de l'effet insignifiantes ou très faibles lors des deux premières prises de mesure (en 5^{ème} année et en 7^{ème} année). Par contre, la troisième mesure (année 9) montre une augmentation importante des performances pour les enfants ayant eu les cours de façon précoce. Cela peut signifier que plus les enfants grandissent, plus les effets des cours d'éveil aux langues reçus tôt dans le cursus scolaire ont des effets sur leurs apprentissages et particulièrement en compréhension orale. Ce point rejoint les propos tenus dans le Pacte pour un enseignement d'Excellence (2020): « La démarche d'Éveil aux langues ne peut porter ses fruits que sur le long terme et dans la continuité » (p. 5). Cependant, nous pourrions interpréter l'augmentation des effets à la croissance naturelle des participants (Hill et al., 2008). Ce cas précis démontre l'importance d'avoir un groupe contrôle permettant de réfuter cette hypothèse.

En tout état de cause, nous ne pouvons baser nos conclusions sur les résultats d'une seule étude, même excellente, dans le cadre d'une revue systématique qui a pour but d'avoir une vision d'ensemble et d'amener des preuves robustes.

4.4.5. Conclusion des études impliquant l'éveil aux langues

Même si la plupart des études sont de qualité supérieure, dans cette partie du rapport, les résultats recueillis sont en nombre insuffisants pour pouvoir fusionner des données présentant des caractéristiques similaires. En outre, il est difficile de comparer des résultats quand les dispositifs et les évaluations mis en place varient considérablement et à tous niveaux. Des pays foncièrement différents sont représentés dans ces études (Éthiopie, Allemagne, Kenya, Suisse, Allemagne, Portugal) et la relation et la motivation des participants envers les langues étrangères sont, sans aucun doute, très diversifiées. Buser (2018) confirme qu'il est complexe de catégoriser des enfants en développement, car ils acquièrent les langues de différentes manières, dans des contextes variés et avec des niveaux de compétence variables dans chaque langue de scolarisation.

Par conséquent, il serait hasardeux de généraliser les résultats obtenus dans ce rapport à l'ensemble de la population ayant des cours d'éveil aux langues. Pourtant, contrairement aux autres pratiques présentées dans cette revue, la Fédération Wallonie-Bruxelles impose à tous les enseignants de maternelle, de première et deuxième primaires, d'enseigner l'éveil aux langues, à tous les élèves, à raison d'une heure par semaine...

5. Limites et perspectives

5.1. Limites :

Idéalement, une revue systématique doit être réalisée par plusieurs personnes, chacun pouvant amener un niveau d'expertise différent (Bibliothèque ULaval, s. d.). Dans le cadre d'un mémoire individuel, cette condition ne peut être respectée, amenant, inévitablement, une certaine subjectivité et un temps de travail plus restreint.

Dans les banques de données, lors de la sélection de thésaurus et du vocabulaire libre, il est possible que certains termes pertinents aient été omis, ne permettant pas d'accéder à des recherches respectant les critères d'éligibilité.

Lors de la sélection des études, plusieurs d'entre elles n'étaient pas accessibles. Malgré une prise de contact via le site ReasearchGate, certains auteurs n'ont jamais répondu. Par conséquent, certains programmes potentiellement efficaces n'ont pas été analysés.

Cette étude a centré ses recherches sur les performances des élèves. Cependant, une pléthore de facteurs peuvent influencer les performances : la motivation, le climat de la classe, le programme et la méthode d'enseignement, l'enseignant, le nombre d'élèves, l'environnement, etc. (Laureline, 2024). Nos recherches n'ont pas tenu compte de l'effet de ces éléments sur les performances et il est probable que certaines études intéressantes n'aient pas été sélectionnées. De plus, nous avons centré nos recherches sur les élèves de l'enseignement primaire ordinaire et, par conséquent, exclu des études analysant les performances d'enfants scolarisés dans l'enseignement spécialisé. Néanmoins, certaines études pouvaient peut-être amener des informations nécessaires à notre revue.

5.2. Perspectives :

Comme nous l'expliquions dans notre cadre théorique : « Il est primordial que l'objectif premier de l'innovation en éducation soit la réussite de tous les élèves (CSÉ, 2006). Alors pourquoi l'impact des recherches en éducation est-il si faible ? L'influence des habitudes, de la politique, des finances sont, selon Slavin (2020), plus tenaces que les preuves. Plusieurs études démontrent que les chefs d'établissement et les enseignants prêtent peu d'importance à la recherche. Certains préfèrent demander conseil à des proches, à des collègues,... plutôt que de regarder les données probantes (Honig & Coburn, 2008 ; Morrison, Ross, Corcoran & Reid, 2014 cités par Slavin 2020). Pourtant, les lacunes de notre système éducatif sont, à l'heure actuelle, connues et reconnues. La preuve d'efficacité est un critère primordial qui permettrait aux programmes d'être munis d'une base solide susceptible d'obtenir de meilleurs résultats pour les enfants. Cette démarche mettrait l'éducation dans une boucle de rétroaction positive

d'innovation, d'évaluation et d'amélioration graduel des programmes (Salvin, 2020) dont elle semble avoir besoin.

6. Conclusion

Notre question de recherche était : « *Les pratiques innovantes proposées dans les formations de la Fédération Wallonie Bruxelles ont-elles un effet sur les performances des élèves de l'enseignement primaire comparativement à ceux évoluant dans des classes « traditionnelles ?* ». Au regard des recherches rigoureuses effectuées dans cette revue systématique, nous ne sommes pas en mesure d'y répondre. En effet, les études sélectionnées ne donnent pas suffisamment de preuves pour tirer une conclusion sur les quatre pratiques innovantes proposées dans les formations de la Fédération Wallonie-Bruxelles, certaines interventions pourraient être bénéfiques mais davantage de preuves sont nécessaires. En effet, nous avons constaté :

- des effets positifs des pratiques sélectionnées (école du dehors, éveil aux langues) mais isolés dans une seule et même étude, ne permettant pas de réaliser une synthèse de la littérature ;
- des effets ne pouvant être combinés par manque de similitude (dans les programmes, dans les outils de récolte de données) ;
- des effets faibles ou insignifiants.

Pour rappel, l'éveil aux langues est un cours devenu obligatoire en FW-B pour les enseignants de maternelle et de première/deuxième primaires. Nous n'avons pourtant pas récolté de preuves solides attestant de son efficacité...

Cependant, les constats de ce rapport ne signifient pas forcément que les pratiques innovantes dont il est question ne sont pas valables : « Le peu d'efficacité que démontrent les résultats d'un programme peut être causé non pas par un manque d'efficacité du traitement et des interventions, mais par des faiblesses dans sa mise en œuvre » (Savignac & Dunbar, 2014, cité par Bissonnette, 2021, p.3).

Efficacité prouvée ou non, les médias parlent abondamment de ces pratiques « à la mode » et un grand nombre d'enseignants choisissent de les utiliser dans leur classe. Des recherches futures seraient plus que nécessaires afin de permettre aux professionnels qui souhaitent baser leurs pratiques sur des preuves, de s'orienter dans la bonne direction. De plus, il serait urgent que les chercheurs et les enseignants collaborent plus abondamment. Si les dispositifs utilisés ont des hauts niveaux de preuves, des changements et des améliorations sont possibles en éducation (Bissonnette, 2020).

7. Bibliographie

- Abbadia, J. (2023, 7 mars). Quelles sont les limites de la recherche et comment les rédiger ? *Mind the Graph Blog*. <https://mindthegraph.com/blog/fr/limites-de-la-recherche/>
- Acheroy, C., Leterme, C., Faniel, A. (2020). Apprendre dehors. Enjeux des pratiques éducatives ancrées dans le milieu. Étude. *Centre d'Expertise et de Ressources pour l'Enfance*. https://www.cere-asbl.be/wp-content/uploads/2021/11/etude_2020_pour_site_2.pdf
- Agegnehu, A. Z., Bachore, M. M., & Ayele, Z. A. (2023). Effects of an explicit rime-based phonics method on the phonological awareness outcomes of Ethiopian EFL children. *Cogent Education*, 10(1). <https://doi.org/10.1080/2331186x.2023.2203582>
- Allam, M.-C., & Chartier, M. (2023). Modifier l'aménagement de la classe pour le bien-être des élèves : vers une forme scolaire flexible ? *Raisons éducatives*, 27, 131–151. <https://doi.org/10.3917/raised.027.0131>
- Alwattar, N. Y., & E.M. Al-Balhan, E. M. (2018). The effectiveness of the e-mind mapping strategy for Sixth-Grade students' achievement level in learning Arabic vocabulary in Kuwait. *The New Educational Review*, 116-125. <https://doi.org/10.15804/tner.2018.53.3.10>
- Baye, A. et Bluge V. (2016). L'éducation fondée sur des preuves. *Rapport réalisé dans le cadre du Pacte d'excellence*. https://orbi.uliege.be/bitstream/2268/252129/1/Baye_Bluge_Education%20fondée%20sur%20des%20preuves_1_10_16.pdf
- Baye, A., Dachtel, D. (2022) Evidence Based education, notes de cours, ecampus de l'université de Liège
- Biais cognitif. (2024, 12 janvier). Perception sélective - Définition du biais cognitif, explication, exemples. *Biais Cognitif*. <https://biais-cognitif.com/biais/perception-selective/>
- Bissonnette, S.; Gauthier, C., & Bocquillon, M. (2020). Pour révolutionner la formation à l'enseignement : proposer des interventions fondées sur des données probantes. *Educação em Perspectiva (Perspective éducation)*, 11, 1-9. <https://periodicos.ufv.br/educacaoemperspectiva/article/view/9010/5828>
- Bissonnette, S. & Boyer, C. (2021). Pour favoriser une implantation réussie des programmes fondés sur des données probantes : prendre en compte les éléments clés d'une mise en oeuvre de qualité. *Formation et profession*, 29(3), 1–6. <https://doi.org/10.18162/fp.2021.a238>
- Boucher, C., & Bouffard, G. (2015). *L'enseignement explicite peut-il convenir au collégial ? : présentation d'une méthode d'enseignement qui a fait ses preuves*. <https://eduq.info/xmlui/handle/11515/34351>

- Budel, S. M. (2020). L'autonomie professionnelle chez l'enseignant innovant. *Thot Cursus*. <https://cursus.edu/fr/21935/lautonomie-professionnelle-chez-lenseignant-innovant>
- Buser, M., & Melfi, G. (2019). Oral proficiency development of K-4 learners of the Swiss two-way immersion program FiBi (Filière Bilingue) in a highly multicultural context. *International Journal Of Bilingual Education And Bilingualism*, 25(1), 328-341. <https://doi.org/10.1080/13670050.2019.1675583>
- Cabelhoule. (2023). Au-delà des données probantes : l'importance de la diversité des connaissances issues des recherches en éducation. *CTREQ - RIRE*. <https://rire.ctreq.qc.ca/au-dela-des-donnees-probantes-limportance-de-la-diversite-des-connaissances-issues-des-recherches-en-education/>
- Charitonidou, A., & Ioannitou, G. (2012). L'autonomie des enseignants: quels éléments caractérisent l'enseignant autonome et comment ils influent sur la décision de la mise en oeuvre d'une innovation pédagogique?. *Synergies France*, (9), 51-59.
- Cheung, A.C., & Slavin, R.E. (2016). How Methodological Features Affect Effect Sizes in Education. *Educational Researcher*, 45, 283 - 292.
- CECP - Bzzz Graphic Design. (s. d.). Conseil de l'Enseignement des Communes et des Provinces asbl » Formation. *Conseil de l'Enseignement des Communes et des Provinces asbl CECP*. <https://www.cecp.be/formation/#fpc>
- Cendros Araujo, R., Gadanidis, G. Cartographie mentale collaborative en ligne dans un programme de formation des enseignants de mathématiques : une étude sur l'interaction des étudiants et la construction des connaissances. *Enseignement mathématique ZDM* 52 , 943-958 (2020). <https://doi.org/10.1007/s11858-019-01125-w>
- Centre for Reviews and Dissemination. (2008). Core principles and methods for conducting a systematic review of health interventions. In *Systematic reviews: CDR's guidance for undertaking reviews in health care* (pp. 3-108). York Publishing Services.
- Clerquin, S., Sastourné-Arrey, A. (2019). L'école dehors : étude de la situation pédagogique “ mon moment à moi ”. Education. 2022. Évaluer une pratique innovante : essai réflexif autour de l'expression théâtrale. *Pédagogiques (Montréal)*, 35(2). <https://doi.org/10.4000/ripes.2116>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^{ème} éd). Lawrence Erlbaum Associates.
- Cortisse, S. (2021, 12 juillet). Enseignement en « classe flexible » : *Cortisse.be*. <https://www.cortisse.be/l/enseignement-en-classe-flexible/>
- Cros, F., & Broussal, D. (2020). Changement et innovation en éducation : deux notions en résonance. *Éducation et socialisation*, 55. <https://doi.org/10.4000/edso.8911>
- Cse, É. (2021). Le dialogue entre la recherche et la pratique en éducation : une clé pour la réussite, *Rapport annuel sur l'état et les besoins de l'éducation 2004-2005 - CSE*.

<https://www.cse.gouv.qc.ca/publications/dialogue-recherche-pratique-rebe-04-05-50-0182/>

- Dachet, D., & Baye, A. (2021). Evidence-Based Education : The (Not So Simple) Case of French-Speaking Belgium. *ECNU review of education*, 4(1), 164-189. <https://doi.org/10.1177/2096531120928086>
- Dambreville S-C. (2014) Les cartes mentales : un changement de paradigme dans les apprentissages. *Communication, Technologies et Développement*, 2014, 1, pp.71-88. file:///Users/carolepeugnieu/Desktop/Cartes_Mentales.pdf
- Davies, P. R. (1999). What is Evidence-based Education ? *British Journal of Educational Studies*, 47(2), 108-121. <https://doi.org/10.1111/1467-8527.00106>
- De Boer, H., Donker, A., & Van Der Werf, G. (2014). Effects of the Attributes of Educational Interventions on Students' Academic Performance. *Review Of Educational Research*, 84(4), 509-545. <https://doi.org/10.3102/0034654314540006>
- De Chambrier, A. & Dierendonck, C. (2022). Vers l'implémentation du modèle de Réponse à l'Intervention dans les systèmes éducatifs d'Europe francophone ?. *L'Année psychologique*, 122, 301-337. <https://doi.org/10.3917/anpsy1.222.0301>
- Degraeve, B. (2022). Chapitre 11. Analyse de puissance et calcul de la taille d'échantillon nécessaire. Dans : , B. Degraeve, *Statistiques en psychologie et neuropsychologie: Méthodes, applications et exercices corrigés* (pp. 349-369). Paris: Dunod.
- Dehors : La nature pour apprendre. (2013). *Symbioses. Le Magazine de L'éducation Relative À L'environnement (ERE)*. <https://www.symbioses.be/pdf/100/Sy-100.pdf>
- De Pietro, J-F. & Matthey, M. (2001). L'éveil aux langues : des outils pour travailler la différence. *Langage & pratiques*, 28, 31-44.
- De Saint-André, M. D., Montésinos-Gelet, I., & Morin, M. (2013). Avantages et limites des approches méthodologiques utilisées pour étudier les pratiques enseignantes. *Nouveaux cahiers de la recherche en éducation*, 13(2), 159-176. <https://doi.org/10.7202/1017288ar>
- Dion, É., Borokhovski, E., Cyr, S., Plante, I. et Charland, P. (2022). Recruter de meilleurs échantillons en éducation. *Didactique*, 3(2), pp. 103-121. <https://doi.org/10.37571/2022.0204>
- Draelants, H., & Dupriez, V. (2015). École : 30 ans de réformes inabouties ? *Société en changement*, 4. UCLouvain. <https://uclouvain.be/fr/decouvrir/presse/actualites/ecole-30-ans-de-reformes-inabouties.html>
- Dulieux, A. (2022, 13 décembre). *10 vertus de la nature sur notre santé physique et mentale*. <https://www.passeportsante.net/>. <https://www.passeportsante.net/fr/Actualites/Dossiers/DossierComplexe.aspx?doc=10-vertus-nature-sur-notre-sante-physique-mentale>

Enquête PISA : les résultats 2018 sont publiés ! (2019, 12 mars). Portail de la Fédération Wallonie-Bruxelles. **Erreur ! Référence de lien hypertexte non valide.**

Enseignement.be – Circulaires 8742. (2022). Enseignement.be.
http://www.enseignement.be/index.php?page=26823&do_id=8997

Enseignement.be - Les recherches en éducation. (s. d.). Enseignement.be.
<http://enseignement.be/index.php?page=24727>

Enseignement.be - L'éveil aux langues - Introduction. (s. d.). Enseignement.be.
[http://www.enseignement.be/index.php?page=24987#:~:text=L'%C3%89veil%20aux%20langues%20vise,voisin%2C%20langues%20europ%C3%A9ennes%E2%80%A6\).](http://www.enseignement.be/index.php?page=24987#:~:text=L'%C3%89veil%20aux%20langues%20vise,voisin%2C%20langues%20europ%C3%A9ennes%E2%80%A6).)

Enseignement.be - Pacte pour un Enseignement d'excellence. (s. d.). Enseignement.be.
<http://www.enseignement.be/index.php?page=28280#:~:text=Le%20Pacte%20pour%20un%20Enseignement,enseignement%20pour%20tous%20les%20%C3%A9l%C3%A8ves.>

Fiske, S. & Taylor, S. (2011). Les stéréotypes. Cognition et biais. Dans : , S. Fiske & S. Taylor (Dir), *Cognition sociale: Des neurones à la culture* (pp. 299-328). Wavre: Mardaga.

Fournier, M. (2016). Un siècle d'éducation nouvelle. Dans : Martine Fournier éd., *Éduquer et Former: Connaissances et débats en Éducation et Formation* (pp. 63-69). Auxerre: Éditions Sciences Humaines. <https://doi.org/10.3917/sh.fourn.2016.01.0063>

Fung, D., Liang, T. The Effectiveness of Collaborative Mind Mapping in Hong Kong Primary Science Classrooms. *Int J of Sci and Math Educ* 21, 899–922 (2023). <https://doi.org/10.1007/s10763-022-10279-1>

Guillon, Q., Hadjikhani, N. & Rogé, B. (2014). L'utilisation de la technique de suivi du regard dans l'étude des troubles du spectre de l'autisme. *L'information psychiatrique*, 90, 827-834. <https://doi.org/10.1684/ipe.2014.1274>

Gokalp, MS, Sharma, M., Johnston, I. et Sharma, M. (2013). Implémentation d'instructions basées sur WebQuest sur la deuxième loi de Newton. *Enseignement des sciences*, 59 (2), 11-19. <https://search.informit.org/doi/10.3316/informit.687316966816099>

Goëta, S., Ho-Pun-Cheung, S. (2022) La standardisation des données ouvertes : favoriser l'interopérabilité, accroître l'impact de l'open data. Observatoire Data Publica. <https://hal.science/hal-03986670/document>

Initiation au mind mapping | *La Ligue de l'Enseignement et de l'Éducation permanente*. (2023). <https://ligue-enseignement.be/formations/initiation-au-mind-mapping>

Initiation au sketchnoting | *Psychoeducation.be*. (s. d.). <https://www.psychoeducation.be/slides/initiation-au-sketchnoting-22>

- Institut inter-réseaux de la Formation professionnelle. (s. d.). https://ifpc.cfwb.be/v5/site_rechform.asp#ALL_ALL_3_0_ALL,FO_0
- Jackel, N., Schurig, M., Van Ackern, I., & Ritter, M. (2022). The impact of early foreign language learning on language proficiency development from middle to high school. *System*, 106, 102763. <https://doi.org/10.1016/j.system.2022.102763>
- Jambon, C. (2022, 27 juin). *7 étapes pour apprendre le sketchnoting (prise de notes visuelle)*. Apprendre, Réviser, Mémoriser. <https://apprendre-reviser-memoriser.fr/apprendre-le-sketchnoting/>
- Jeanlouisestrade. (2019, 6 septembre). *La taille de l'effet*. WordPress.com. <https://kinotesfr.wordpress.com/2019/09/09/la-taille-de-leffet/>
- Kamper S. J. (2019). Interprétation des résultats 2 - Signification statistique et signification clinique : Relier les preuves à la pratique. *The Journal of orthopaedic and sports physical therapy*, 49(7), 559-560.
- Kaplan, A. (1960). *The Conduct of Inquiry: Methodology for Behavioral Science*. San Francisco, CA: Chandler Publishing Company.
- Larose, F., Couturier, Y., Bédard, J. & Charette, S. (2011). Entre discipline et profession : la question des bonnes pratiques guidées par les résultats probants de la recherche (*evidence based practice*) en formation à l'enseignement. *Les Sciences de l'éducation - Pour l'Ère nouvelle*, 44, 31-48. <https://doi.org/10.3917/lsdle.442.0031>
- Laruelle, O. (2009). Les Politiques de formation continuée des enseignants : Une approche comparative (Presse universitaire de Namur). *Collection Politique*. 5-32. <https://books.google.be/books?id=QKLYDROb7hkC&printsec=frontcover&hl=fr#v=onepage&q&f=false>
- L'éducation fondée sur des preuves : définition et enjeux. (2022). *Institut des compétences et de l'innovation*. <https://www.sciencespo.fr/institut-competences-innovation/fr/actualites/l-education-fondee-sur-des-preuves-definition-et-enjeux>
- Lourenço, M., & Andrade, A. I. (2013). Promoting phonological awareness in pre-primary education : possibilities of the 'awakening to languages' approach. *Language Awareness*, 23(4), 304-318. <https://doi.org/10.1080/09658416.2013.783585>
- Maulini, O., & Capitanescu Benetti, A. (2015). Que penser. . . .de la classe flexible ? *Faculté de Psychologie et des Sciences de L'éducation*. <https://www.unige.ch/fapse/SSE/teachers/maulini/publ-2001.pdf>
- McGough, J. J., & Faraone, S. V. (2009). Estimating the size of treatment effects: Moving beyond P values. *Psychiatry (Edgmont (Pa.: Township))*, 6(10), 21-29.
- Mind Map selon Tony Buzan : Règles du mind mapping et exemple illustré - Créativité.Net. (s. d.-b). <https://www.creativite.net/mind-mapping-de-tony-buzan/mind-map.php>

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2), 364-386

Ordre des psychoéducateurs et psychoéducatrices du Québec. (2022, août 29). Ligne directrice sur l'utilisation des instruments de mesure. *Ordre des Psychoéducateurs et Psychoéducatrices du Québec*. https://ordrepsed.qc.ca/wp-content/uploads/2022/03/Lignes_directrices_sur_l_utilisation_des_instruments_de_mesure.pdf

Pacte pour un Enseignement d'Excellence. (s. d.). Portail de la Fédération Wallonie-Bruxelles. https://www.federation-wallonie-bruxelles.be/index.php?id=detail_article&no_cache=1&tx_cfwbarticlefe_cfwbarticlefront%5Baction%5D=show&tx_cfwbarticlefe_cfwbarticlefront%5Bcontroller%5D=Document&tx_cfwbarticlefe_cfwbarticlefront%5Bpublication%5D=664&cHash=c1a49278999ea1e89622402808b04533

Pacte pour un Enseignement d'excellence. Le monde évolue. L'école aussi. (2023, 11 janvier). <https://pactepourunenseignementdexcellence.cfwb.be/>

Page, M., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, 372, Article 71. <https://doi.org/10.1136/bmj.n71>

Pinard, R., Potvin, P. & Rousseau, R. (2004). Le choix d'une approche méthodologique mixte de recherche en éducation. *Recherches qualitatives*, 24, 58-82. <https://doi.org/10.7202/1085563ar>

Porter, A. (2019). An early start to foreign language literacy in English primary school classrooms. *Language Learning Journal*, 48(5), 656-671. <https://doi.org/10.1080/09571736.2019.1632918>

Potvin, P. (2015). Pratiques innovantes en éducation. *Catégorie Éducation*. <https://pierrepotvin.com/wp/index.php/2015/05/20/100/>

Programme P.A.R.L.E.R en FW-B. (2021). https://www.equale.uliege.be/cms/c_8662000/fr/programme-p-a-r-l-e-r-en-fw-b?id=c_8662000

Rajteric, L. (2024, 16 février). Carte mentale : le guide ultime pour booster la créativité et l'organisation [2024] • Asana. *Asana*. <https://asana.com/fr/resources/what-is-mindmap>

Ramus, F. (2019, 8 janvier). Vers une éducation fondée sur des preuves | Ramus méninges. <https://scilogs.fr/ramus-meninges/vers-education-fondee-preuves/>

Revue systématique | Bibliothèque | ULaval. (s. d.). <https://www.bibl.ulaval.ca/services/soutien-a-ledition-savante-et-a-la-recherche/syntheses-de-la-connaissance/realiser-une-revue-systematique>

- Roques, N. (2022). Introduction au calcul de la taille d'effet globale d'une intervention dans les méta-analyses en sciences de l'éducation. *Mesure et évaluation en éducation*, 45(3), 1–32. <https://doi.org/10.7202/1106852ar>
- Sacré, M., Lafontaine, D. & Toczec, M.-C. (2021). Comprendre et concevoir des revues systématiques de la littérature en sciences de l'éducation et de la formation. *Nouveaux cahiers de la recherche en éducation*, 23(2), 1–27. <https://doi.org/10.7202/1085361ar>
- Shaikh, M. (2023, 3 février). Déterminer la taille de l'échantillon : définition, formule et exemple. *QuestionPro*. <https://www.questionpro.com/blog/fr/determining-sample-size/#:~:text=Si%20la%20taille%20de%20votre,entra%C3%A9ne%20des%20r%C3%A9sultats%20non%20concluants>.
- Sham. (2013, août 6). *Le sophisme du petit échantillon* | Sham and Science. <https://blogs.univ-poitiers.fr/n-yeganefar/2013/08/06/le-sophisme-du-petit-echantillon/>
- Slavin, R.E., & Smith, D. (2009). The Relationship Between Sample Sizes and Effect Sizes in Systematic Reviews in Education. *Educational Evaluation and Policy Analysis*, 31, 500 - 506.
- Slavin, R. E. (2020). How evidence-based reform will transform research and practice in education. *Educational Psychologist*, 55(1), 21-31. <https://doi.org/10.1080/00461520.2019.1611432>
- Spw. (2022, 5 mars). Ecole du Dehors : la Wallonie booste la formation des enseignants. Céline TELLIER- Ministre. <https://tellier.wallonie.be/home/presse--actualites/communiques-de-presse/presses/ecole-du-dehors--la-wallonie-booste-la-formation--des-enseignants.html>
- Tisseron, S. (2017). Les dérives de l'empathie. *L'école des parents*, 623, 7-11. <https://doi.org/10.3917/epar.623.0007>
- Statistique Canada, G. (2021, 2 septembre). 3.2.2 Échantillonnage probabiliste. <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch13/prob/5214899-fra.htm>
- Tong, F., Irby, B. J., Lara-Alecio, R., Yoon, M., & Mathes, P. G. (2010). Hispanic English Learners' Responses to Longitudinal English Instructional Intervention and the Effect of Gender : A Multilevel Analysis. *The Elementary School Journal/The Elementary School Journal*, 110(4), 542-566. <https://doi.org/10.1086/651195>
- Tous dehors Belgique. (s. d.). *Tous Dehors Belgique*. <https://tousdehors.be/?PageAccueil>
- Tufanaru, C., Munn, Z., Aromataris, E., Campbell, J., Hopp, L. (2020). Systematic reviews of effectiveness. In: Aromataris, E., Munn, Z. (Eds.), *JBIM Manual for Evidence Synthesis*. <https://doi.org/10.46658/JBIMES-20-04>
- Vista do Para revolucionar a formação docente. (s. d.). <https://periodicos.ufv.br/educacaoemperspectiva/article/view/9010/5828>

- Wagnon, S. (2018). Les pédagogies alternatives en France aujourd'hui : essai de cartographie et de définition. *Tréma*, 50. <https://doi.org/10.4000/trema.4174>
- Wallonie-Bruxelles Enseignement. (s. d.). WBE. <https://www.wbe.be/>
- Wawire, B. A., & Kim, Y. G. (2018). Cross-Language Transfer of Phonological Awareness and Letter Knowledge : Causal Evidence and Nature of Transfer. *Scientific Studies Of Reading*, 22(6), 443-461. <https://doi.org/10.1080/10888438.2018.1474882>
- Yang, T. (2017, 8 juin). La carte mentale, un outil didactique d'aide à la mémorisation pour l'enseignement/apprentissage du lexique. <https://hal.science/hal-02013296>
- Young, S. et Eldermire, E. (2017). The Big Picture: Finding, Evaluating, Applying
- Zorman, M., Bressoux, P., Bianco, M., Lequette, C., Pouget, G., & Pourchet, M. (2015). "PARLER" : a program to prevent learning difficulties. *Revue Française de Pédagogie/Revue Française de PéDagogie*, 193, 57-76. <https://doi.org/10.4000/rfp.4890>

8. Annexes

Annexe 1 : Qualité des études incluses (PRISMA, 2020)

1.A. « L'école du dehors »

1.

Auteurs	Date	Indicateur	Évaluation	Score
Otte, C.R et al.	2019	Design de recherche (expérimental)	1	6/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	1	

2.

Auteurs	Date	Indicateur	Évaluation	Score
Quibell, T.; Charlton, J. ; Loi, J.	2017	Design de recherche (expérimental)	1	4/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

3.

Auteurs	Date	Indicateur	Évaluation	Score
Wells et al.	2015	Design de recherche (expérimental)	1	5/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0 (âge)	

4.

Auteurs	Date	Indicateur	Évaluation	Score
	2014	Design de recherche (expérimental)	1	4/6

Ting, K.L.; Siew, N.M.		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

5.

Auteurs	Date	Indicateur	Évaluation	Score
Fancovicova, J. ; Prokop, P.	2011	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

6.

Auteurs	Date	Indicateur	Évaluation	Score
Chang, C- S. ; Chen, T-S. ; Hsu, W-H.	2011	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	0	

7.

Auteurs	Date	Indicateur	Évaluation	Score
Carrier, S.	(2009)	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0	

1.B. Cartes mentales

1.

Auteurs	Date	Indicateur	Évaluation	Score
Wan, H; Yu, S.	2023	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	0	

2.

Auteurs	Date	Indicateur	Évaluation	Score
Fung, D.; Liang, T.	2023	Design de recherche (expérimental)	1	5/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

3.

Auteurs	Date	Indicateur	Évaluation	Score
Merchie, E.; Catrysse, L.; Van Keer, H.	2021	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

4.

Auteurs	Date	Indicateur	Évaluation	Score
Aliasier, A.M..	2017	Design de recherche (expérimental)	1	3/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

5.

Auteurs	Date	Indicateur	Évaluation	Score
Lan, Y. J., Sung, Y. T., Cheng, C. C., & Chang, K. E.	2015	Design de recherche (expérimental)	1	1/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	0	

6.

Auteurs	Date	Indicateur	Évaluation	Score
Balim, A. G.	2013	Design de recherche (expérimental)	1	3/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

7.

Auteurs	Date	Indicateur	Évaluation	Score
Poland M. ; van Oers, B.	2007	Design de recherche (expérimental)	1	5/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0	

1.C. Classes flexibles

1.

Auteurs	Date	Indicateur	Évaluation	Score
Hardjansyah, F.; Abuyamin, R.; Muhammad M.	2022	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	0	

2.

Auteurs	Date	Indicateur	Évaluation	Score
Stapp, A. (2018)	2018	Design de recherche (expérimental)	1	2/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

1.D. Éveil aux langues

1.

Auteurs	Date	Indicateur	Évaluation	Score
Agegnehu, A.Z. et al.	2023	Design de recherche (expérimental)	1	4/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

2.

Auteurs	Date	Indicateur	Évaluation	Score
Jaekel, N., Schurig, M., Van Ackern, I., & Ritter, M.	2022	Design de recherche (expérimental)	1	6/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	1	

3.

Auteurs	Date	Indicateur	Évaluation	Score
Wawire, B. A., & Kim, Y. G.	2018	Design de recherche (expérimental)	1	5/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	0	
		Équivalence au prétest (<25% d'écart-type)	1	

4.

Auteurs	Date	Indicateur	Évaluation	Score
Buser, M.	2020	Design de recherche (expérimental)	1	3/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	0	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0	

5.

Auteurs	Date	Indicateur	Évaluation	Score
Maluch, J. T., Neumann, M., & Kempert, S.	2016	Design de recherche (expérimental)	1	4/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	1	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0	

6.

Auteurs	Date	Indicateur	Évaluation	Score
Lourenço, M., & Andrade, A. I.	2013	Design de recherche (expérimental)	1	4/6
		Taille des effectifs (>30 sujets par groupe)	0	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	1	

7.

Auteurs	Date	Indicateur	Évaluation	Score
Porter, A. (2019)	2019	Design de recherche (expérimental)	0	3/6
		Taille des effectifs (>30 sujets par groupe)	1	
		Outil de récolte de données (indépendant)	1	
		Outil de récolte de données (standardisé)	0	
		Durée de l'étude (>12 semaines)	1	
		Équivalence au prétest (<25% d'écart-type)	0	

Annexe 2 : Programmes détaillés des études sélectionnées

2.A. Programme détaillé : « École du dehors »

Étude	Programme	Intervention dans le groupe contrôle
1	L'éducation hors de la classe (EOTC) de 4,7h/semaine en moyenne durant un an. Les enseignants choisissaient les activités et devaient fournir des rapports tout au long de l'année sur leur utilisation hebdomadaire de l'EOTC, les matières enseignées et le cadre dans lequel l'enseignement avait lieu avec les classes concernées.	L'éducation hors de la classe (EOTC) de 0,9h/semaine en moyenne.
2	Wilderness Schooling (WS) est un programme manuel de plans de cours et de ressources d'activités destiné à guider un programme d'enseignement en extérieur d'une durée de six jours. Chaque programme WS est dispensé par deux Wilderness Practitioners (WPs), dont l'un est un enseignant qualifié du primaire ou du secondaire et l'autre un assistant.	Néant
3	20 leçons en extérieur une fois par semaine contenant des cours de jardinage axés sur la nutrition, l'horticulture et la science végétale. Le programme a été conçu par une équipe de spécialistes sur base de 17 programmes existants fondés sur des données probantes (source manquante).	Néant
4	Développement des compétences scientifiques d'observation, de prédiction et de déduction tout au long des leçons dans la cour de récréation.	Néant
5	Plantation d'arbres + exposé sur les fonctions de l'économie forestière et sur la plantation d'agrément + enseignement non formel de la biologie axé sur la botanique avec trois conférences de 120 minutes + visité d'une prairie adjacente.	Aucune activité hormis la visite d'une prairie adjacente.
6	Utilisation du programme WebQuest dans un environnement éducatif en plein air + activités pédagogiques dans la cour de l'école primaire. Le Webquest est une activité de recherche structurée en ligne réalisée par l'élève (Gokalp et al., 2013).	GC1:enseignement traditionnel en classe GC2 :enseignement traditionnel en classe combiné avec WebQuest GE ; enseignement en extérieur combiné avec WebQuest.

7	Activités liées aux connaissances des concepts scientifiques/environnementaux comprenant les habitats, l'adaptation, l'interdépendance, la classification et les organismes de la cour d'école.	Non communiqué
---	---	----------------

2.B. Programme détaillé de l'utilisation des cartes mentales

Étude	Programme	Intervention dans le groupe contrôle
1.	Utilisation d'un système d'apprentissage en ligne et d'une carte cognitive d'apprentissage. L'enseignant a construit le graphique « comment utiliser le logiciel Excel » et les chercheurs l'ont importé dans le système d'apprentissage en ligne. L'enseignante a réalisé les activités d'apprentissage. Ensuite, les chercheurs ont donné aux deux groupes une formation adéquate sur l'utilisation du système d'apprentissage en ligne proposé. La fonction de carte cognitive d'apprentissage a également été présentée au groupe expérimental (Wan, 2023).	Utilisation d'un système d'apprentissage en ligne sans carte cognitive.
2.	Les effets de la cartographie mentale collaborative en 20 séances. T1 : groupe témoin ; T2 : collaboration ; T3 : collaboration + carte mentale Le groupe expérimental (T 3) a intégré à la fois le principe pédagogique du travail de groupe et l'utilisation efficace des cartes mentales dans l'enseignement des sciences (Fung, 2023).	Tous les étudiants ont suivi une formation sur la manière de refléter leurs idées à travers la construction de cartes mentales.
3.	Utilisation de cartes mentales pour la lecture d'un texte informatif de 316 mots et 35 phrases sur le réchauffement climatique. Le texte global et les paragraphes individuels du texte ont été analysés par l'Institut néerlandais pour le développement de tests (Centraal instituut voor toetsontwikkeling ; Cito). Une carte mentale d'expert dessinée par ordinateur, construite par un chercheur qualifié, a été incluse dans l'expérience (Merchie, 2021). Le groupe MMT étudie de la carte mentale avant de réaliser l'apprentissage du texte Le groupe TMM : étudie du texte avant l'apprentissage de la carte mentale	Étude d'un texte informatif de trois paragraphes sur le réchauffement climatique sans carte mentale.
4.	Utilisation d'une carte mentale électronique. L'enseignant a reçu (1) un guide avec les objectifs et la méthode d'enseignement appropriée. (2) Un	Enseignement traditionnel

	aperçu de la stratégie d'enseignement par cartes électroniques (3) Le plan d'enseignement (4) Des outils et supports pédagogiques. Le guide de l'enseignant a été soumis à un groupe de réviseurs spécialisés dans les programmes (Aliaser, 2017). Ensuite, l'unité d'étude proposée a été enseignée au groupe expérimental à l'aide de cartes mentales électroniques.	
5.	Le programme comprend différentes stratégies de préécriture coopérative assistée par ordinateur (brainstorming basé sur le texte (BS), dessin (DR) et cartographie mentale (MM). Trois kits pédagogiques conçus par les auteurs ont été utilisés. Chaque kit comprenait du matériel pédagogique et des instructions détaillées.	Les élèves des trois groupes ont reçu un enseignement identique de l'écriture en anglais
6	Technique de carte mentale manuscrite en groupe Le programme comprenait 5 expériences ouvertes et des feuilles de travail que les élèves pouvaient étudier avec leurs amis pendant les expériences ou les activités + cartes cognitives à compléter. Pas d'informations supplémentaires concernant les cartes mentales.	Méthodes pédagogiques traditionnelles

2.C. Programme détaillé des classes flexibles

Étude	Programme	Intervention dans le groupe contrôle
1	Les variations du style de disposition des sièges utilisées comprennent le face à face, le chevron, le groupe, le séminaire et la conférence. Pas d'autres informations	Disposition traditionnelle des sièges
2	Utilisation de sièges alternatifs, sous la forme de coussins en forme de disque pour observer les élèves à la tâche et hors tâche durant un cours de mathématique.	Pas de groupe contrôle

2.D. Éveil aux langues

Étude	Programme	Intervention dans le groupe contrôle
1	Intervention de lecture utilisant une méthode phonétique explicite basée sur la rime. 36 leçons d'anglais de 40 minutes ont été dispensés par des professeurs d'anglais non natifs (extérieur à l'école)	Néant
2	Le groupe expérimental représente des débutants précoces : cours de langue étrangère en 1 ^{ère} primaire. L'étude faisait partie d'un projet multidisciplinaire et multi-écoles (<i>Ganz In - All-Day Schools for a Brighter Futur</i>).	Le groupe contrôle correspond à des élèves débutants tardifs: cours de langue étrangère en 3 ^{ème} primaire
	La différence d'exposition aux cours de langues étrangères entre les deux groupes était de 105 cours de 45 minutes, pour un total de 78 ³ / ₄ heures	
3	Formation de 8 semaines sur la conscience phonologique et la connaissance des lettres en kiswahili et en anglais. Les enfants étaient retirés de la classe pour chaque session. L'enseignement s'occupait de groupes de trois enfants pendant 20 minutes, trois fois par semaine. Les principes fondés sur la recherche pour un enseignement efficace de la conscience phonologique ont été utilisés adaptant des jeux métalinguistiques et des exercices d'évaluation et d'enseignement de la conscience phonologique (Lane & Pullen, Citation2004)	Les enfants du groupe contrôle ont continué à suivre le programme régulier de Tusome pour apprendre quotidiennement le kiswahili et l'anglais pendant 30 minutes dans chaque langue chaque semaine.
4	Programme d'immersion bidirectionnelle de l'école publique suisse Filière Bilingue (FiBi). Le FiBi est une alternative éducative basée sur le choix qui intègre les apprenants francophones et germanophones et prône le bilinguisme et la réussite scolaire. Chaque classe FiBi est composée d'un tiers d'apprenants allophones n'ayant ni le français, ni l'allemand ou le suisse allemand comme première langue. Le temps d'enseignement est égal en allemand (suisse) et en français à tous les niveaux	Néant
5	Il n'y a pas de programme spécifique mais simplement l'effet du bilinguisme ou du monolinguisme à la maison. La définition des élèves issus de minorités linguistiques utilisée dans la présente étude inclut tous les enfants qui ont été exposés à deux langues dans la petite enfance (Bialystok, 2001).	Enfants monolingue
6	Un programme d'éveil aux langues (AtL) sur le développement de la conscience phonologique (AP). Les activités comprenaient de la rédaction linguistique de l'individu et du groupe, l'écoute de chansons dans	Les enfants du groupe témoin n'étaient pas présents dans les activités AtL et n'ont été soumis

	différentes langues, l'apprentissage des traditions de Pâques dans d'autres pays, l'écriture et la comparaison de mots en hiéroglyphes grecs, mandarins et égyptiens et l'apprentissage des noms d'animaux en voie de disparition en cherokee et en inuktitut.	qu'aux deux séries de tests PA.
--	--	---------------------------------

Annexe 3 : Outils de récolte de données des études sélectionnées

3.A. « École du dehors »

Études	Outil de récolte de données
1	<p>Test de lecture standardisé de Hogrefe (Moller, 2012) appelé « test de lecture de phrases textuelles » reconnu comme ayant une bonne validité et fiabilité.</p> <p>Cette étude fait partie de l'étude danoise TEACHOUT, une étude d'évaluation interdisciplinaire à méthodes mixtes qui vise à comprendre comment une exposition régulière à l'EOTC (éducation hors de la classe) influence l'activité physique, les relations sociales, la motivation scolaire, le bien-être et l'apprentissage chez les enfants fréquentant la troisième année. à six ans (8 à 13 ans) dans les écoles publiques danoises (Nielsen et al., 2016, cités par Otte et al., 2019).</p>
2	<p>-Tests en anglais, lecture, écriture et mathématiques réalisés en 3 temps (T1 → pré-test ;T2 ;T3).</p> <p>-Tests T2 et T3 donnés par les enseignants et transmission des données par mail aux chercheurs.</p> <p>-Pas d'informations supplémentaires</p>
3	<p>Test standardisé de connaissances scientifiques via un questionnaire à choix multiples de 7 éléments sélectionnés dans l'enquête « Eating from the Garden Curriculum » de l'Université du Missouri. La validité du contenu de l'instrument a été établie par l'équipe d'élaboration du programme Healthy Gardens, Healthy Youth.</p>
4	<p>Deux tests équivalents de compétences en processus scientifique, chacun composé de 20 questions à choix multiples d'observation, de classification, d'inférence, de prédiction, de communication et de contrôle des variables.</p> <p>-Pas d'informations supplémentaires</p>
5	<p>- Le test comprenait 13 items composés de questions ouvertes, de questions à choix multiples, des questions additives pour compléter un schéma de réseau alimentaire et une question relative au dessin de l'écosystème d'une prairie.</p> <p>-Pas d'informations supplémentaires</p>

6	Pré-test et post test non équivalents créés par les enseignants participants. Un questionnaire a été conçu sur la base de l'évaluation de la satisfaction des apprenants fournie le chercheur. Le questionnaire utilisait une échelle de Likert en 5 points.
7	Instrument conçu par le chercheur basé sur l'échelle CHEAKS (Children's Environmental Attitudes and Knowledge Scale) (Leeming, Dwyer, & Bracken, 1995)

3.B. Cartes mentales

Études	Outil de récolte de données
1	L'enseignant a créé l'évaluation. Parallèlement à l'interaction des étudiants, le système proposé pouvait automatiquement évaluer le degré de maîtrise d'une certaine compétence pour chaque étudiant. Chaque étudiant des deux groupes a soumis un tableur électronique Excel, qui a été évalué par l'enseignant du cours de manière anonyme.
2	Les étudiants ont reçu un pré-test et post-test identique adapté de TIMSS pour les évaluer avant et après l'intervention pédagogique. Le test comprenait 30 items. De plus, les cartes mentales des étudiants ont été collectées tout au long de l'intervention, avec plus de 200 collectes à la fin de la période de recherche. Les cartes ont ensuite été soumises à une analyse de contenu pour révéler l'impact de la cartographie mentale collaborative sur la compréhension scientifique des élèves (Fung, 2023).
3	Pour évaluer la compréhension et la mémoire des élèves, une mesure post-test à plusieurs niveaux a été administrée (Kim et al., 2018) composée de trois sous-tests à réponse verbale + questions à choix multiples + 5 questions de déduction. Les étudiants dans les conditions TMM et MMT n'ont reçu aucune information supplémentaire sur la façon de lire ou d'interpréter la carte mentale, que ce soit avant, pendant ou après
4	10 questions sur les règles grammaticales et 8 questions sur la grammaire. Pas d'informations supplémentaires.
5	Les performances rédactionnelles des participants ont été évaluées à l'aide de deux tests conçus par les auteurs (Lan et al., 2012) : le Grammaticality Judgment Test (GJT) et l'English Composition Test (ECT). Les deux tests ont été revus par trois experts de l'enseignement de l'écriture en langue étrangère et administrés à 165 élèves de cinquième année de la même école dans le cadre d'une étude pilote
6	Test de réussite scolaire en sciences en plusieurs questions créé par les chercheurs et vérifié par des experts. Pas d'informations supplémentaires.

3.C. Classe flexible

Études	Outil de récolte de données
1	Les données ont été recueillies à l'aide d'un test pour collecter des informations sur les résultats d'apprentissage des élèves après avoir participé à l'apprentissage pour répondre à des problèmes, à savoir augmenter la motivation d'apprentissage des élèves en changeant de place (Hardiansyah et al., 2022).
2	Les données ont été collectées pendant neuf périodes d'observation de trente minutes par trois codeurs. Ils ont tous reçu une formation avant l'étude, qui consistait à coder des séances enregistrées sur vidéo et en direct. La fiabilité inter-évaluateurs a été déterminée avant l'étude et alignée sur les études de recherche antérieure. Pour cette étude, les comportements ont été codés comme étant soit à la tâche, soit hors tâche.

3.D. Éveil aux langues

Études	Outil de récolte de données
1	Un test a été développé sur la base du Phonological Awareness Test (PAT) de Robertson et Salter (1997). La fiabilité des tests a été vérifiée à l'aide d'un contrôle de fiabilité test-retest avant qu'ils ne soient utilisés pour collecter les données.
2	Les données sur les étudiants LS ont été collectées au cours des années 5, 7 et 9. Les élèves n'ont été inclus dans les analyses que s'ils avaient participé à toutes les évaluations. Tout au long du projet, les deux groupes ont reçu des évaluations linguistiques et cognitives identiques. Les tests de compréhension orale et écrite étaient standardisés et étaient auparavant utilisés et validés dans des évaluations à grande échelle en Allemagne.
3	Les participants ont été testés sur neuf tests en kiswahili et en anglais. Ils ont été donnés en 2 sessions distinctes, espacées de 2 à 3 jours. Chaque session durait environ 20 à 30 minutes. Chaque enfant a été testé individuellement. Toutes les évaluations ont été administrées à l'aide d'un outil d'évaluation sur tablette, le Tangerine-TM par le premier auteur et cinq assistants de recherche extérieur à l'école et formés. Ils corrigeaient « à l'aveugle » L'Early Grade Reading Assessment (EGRA), qui inclut les compétences en lecture des enfants des premières années d'études a été utilisé mais sans chronométrer comme dans la version initiale.
4	La méthode choisie est un entretien semi-structuré, orienté performance dans leurs deux langues de scolarisation. Chaque entretien durait entre 20 et 30 minutes. Les données permettant de mesurer les compétences en langue orale ont été collectées chaque année pour les trois premiers groupes FiBi et de la classe témoin.
5	Cette étude examine l'effet du bilinguisme sur la réussite en anglais comme langue étrangère du primaire au secondaire. Les résultats en anglais ont été évalués à l'aide d'un test de Cloze, stylo et papier, en sixième et huitième années. Cet instrument a

	été développé non seulement pour cette étude, mais a également été mis en œuvre dans d'autres évaluations à grande échelle en allemand.
6	À l'aide de méthodes de recherche mixtes, des données ont été recueillies à partir d'enregistrements vidéo de sept sessions AtL et de tests PA pour le groupe expérimental et le groupe témoin. Le test comprenait six tâches différentes, visant à évaluer la conscience des enfants des syllabes et des phonèmes (l'identification de la première syllabe, l'identification du premier phonème, la suppression de la première syllabe, la suppression du premier phonème, la segmentation des syllabes et la segmentation des phonèmes) Ils comprenaient 14 items chacun, précédés de deux items d'entraînement.

Annexe 4 : Équations de recherche

4.A. École du dehors

Intervention	Design de recherche	Population
“outside school” or “outside learning” or “forest school” or “outdoor learning” or “outside activit*” or “outside education” or “education outside the classroom” or “outdoor education” or “outdoor teaching” or “wood school” or “contact with nature” or “interaction with nature” or “Wilderness schooling” or “learning in natural environments” or “outside the classroom” or “access to green space”	ABA design*.mp or ABAB design*.mp. or Academic Achievement/ or Achievement Gains/ or Best practices/ or Causal.mp. or Comparative Analysis/ or Comparison group*.mp or Control groups/ or Effect Size/ or Effectiv*.mp. or Evidence/ or Experimental.mp. or Experimental Groups/ or Experimental Programs/ or Experimental Schools/ or experimentation.mp. or impact.mp. or Matched Groups/ or Meta Analysis/ or Quasi experimental.mp. or Quasiexperimental Design/ or RCT.mp. or Randomized Controlled Trials/ or Regression Discontinuity.mp. or Systematic review.mp. or “natural experiment” .mp. or matched.mp.	Elementary Schools/ or Elementary School Curriculum/ or Elementary School Students/ or Elementary Education/ or Elementary School Science/ or Elementary Secondary Education/ or Elementary School Mathematics/ or Elementary School Teachers/ or Grade 1/ or Grade 2/ or Grade 3/ or Grade 4 /or Grade 5/ or Grade 6/

4.B. Les cartes mentales

Intervention	Design de recherche	Population
“mind mapping” or “mind map” or “mindmap” or “mental map” or “mental cartography” or “mental mapping” or “mental schema” or “sketchnoting” or “sketchnot*” or “illustrated recording” or “doodle note” or “heuristic map” or “heuristic card” or “mental imagery” or “cognitive map*” or “schemata”	ABA design*.mp or ABAB design*.mp. or Academic Achievement/ or Achievement Gains/ or Best practices/ or Causal.mp. or Comparative Analysis/ or Comparison group*.mp or Control groups/ or Effect Size/ or Effectiv*.mp. or Evidence/ or Experimental.mp. or Experimental Groups/ or Experimental Programs/ or Experimental Schools/ or experimentation.mp. or impact.mp. or Matched Groups/ or Meta Analysis/ or Quasi experimental.mp. or Quasiexperimental Design/ or RCT.mp. or Randomized Controlled Trials/ or Regression Discontinuity.mp. or Systematic review.mp. or “natural experiment” .mp. or matched.mp.	Elementary Schools/ or Elementary School Curriculum/ or Elementary School Students/ or Elementary Education/ or Elementary School Science/ or Elementary Secondary Education/ or Elementary School Mathematics/ or Elementary School Teachers/ or Grade 1/ or Grade 2/ or Grade 3/ or Grade 4 /or Grade 5/ or Grade 6/

4.C. Les classes flexibles

Intervention	Design de recherche	Population
“dynamic seating” or “flexible learning” or “flexible classroom” or “flexible teach*” or “Flexible Seat*” or “Classroom design” or “special organization” or “Class Organization” or “innovative learning environment” or “Space Utilization” or “Furniture”	ABA design*.mp or ABAB design*.mp. or Academic Achievement/ or Achievement Gains/ or Best practices/ or Causal.mp. or Comparative Analysis/ or Comparison group*.mp or Control groups/ or Effect Size/ or Effectiv*.mp. or Evidence/ or Experimental.mp. or Experimental Groups/ or Experimental Programs/ or Experimental Schools/ or experimentation.mp. or impact.mp. or Matched Groups/ or Meta Analysis/ or Quasi experimental.mp. or Quasiexperimental Design/ or RCT.mp. or Randomized Controlled Trials/ or Regression Discontinuity.mp. or Systematic review.mp. or “natural experiment” .mp. or matched.mp.	Elementary Schools/ or Elementary School Curriculum/ or Elementary School Students/ or Elementary Education/ or Elementary School Science/ or Elementary Secondary Education/ or Elementary School Mathematics/ or Elementary School Teachers/ or Grade 1/ or Grade 2/ or Grade 3/ or Grade 4 /or Grade 5/ or Grade 6/

4.D. L'éveil aux langues

Intervention	Design de recherche	Population
“Early language learning” or “early learning of foreign languages” or “early language awareness” or “early teaching of foreign languages” or “prevention of foreign languages” or “awareness of language” or “awareness of foreign languages” or “language awareness” or “foreign-languages awareness” or “diversity of languages in elementary school” or “awakening to languages” or “foreign linguistics in elementary school”	ABA design*.mp or ABAB design*.mp. or Academic Achievement/ or Achievement Gains/ or Best practices/ or Causal.mp. or Comparative Analysis/ or Comparison group*.mp or Control groups/ or Effect Size/ or Effectiv*.mp. or Evidence/ or Experimental.mp. or Experimental Groups/ or Experimental Programs/ or Experimental Schools/ or experimentation.mp. or impact.mp. or Matched Groups/ or Meta Analysis/ or Quasi experimental.mp. or Quasiexperimental Design/ or RCT.mp. or Randomized Controlled Trials/ or Regression Discontinuity.mp. or Systematic review.mp. or “natural experiment” .mp. or matched.mp.	Elementary Schools/ or Elementary School Curriculum/ or Elementary School Students/ or Elementary Education/ or Elementary School Science/ or Elementary Secondary Education/ or Elementary School Mathematics/ or Elementary School Teachers/ or Grade 1/ or Grade 2/ or Grade 3/ or Grade 4 /or Grade 5/ or Grade 6/