

Comprendre et concevoir des revues systématiques de la littérature en sciences de l'éducation et de la formation

Understanding and designing systematic literature reviews in education and training

Comprender y diseñar revisiones sistemáticas de literatura en ciencias de la educación y la formación

Margault Sacré, Dominique Lafontaine and Marie-Christine Toczek

Volume 23, Number 2, 2021

URI: <https://id.erudit.org/iderudit/1085361ar>

DOI: <https://doi.org/10.7202/1085361ar>

[See table of contents](#)

Publisher(s)

Faculté d'éducation, Université de Sherbrooke

ISSN

1911-8805 (digital)

[Explore this journal](#)

Cite this article

Sacré, M., Lafontaine, D. & Toczek, 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>

Article abstract

Systematic literature reviews offer a synthesis of scholarly knowledge on a given research subject. The transparency of the methodology is a benchmark of the reliability of the review and the proposed conclusions. Evidence-based education defends this methodology, but generally ignores non-experimental studies. The purpose of this article is to help understand and design systematic reviews in education and training, including different types of research, from experimental studies to comprehensive and exploratory studies. The article details the steps involved in the construction of a systematic review and proposes a writing script for any researcher engaging in writing a systematic review, irrespective of their level of expertise.

nouveaux cahiers

NERÉ

de la recherche en éducation

VOLUME 23
n° 2 | 2021



Dans ce numéro...

Comprendre et concevoir des revues systématiques de la littérature en sciences de l'éducation et de la formation

Margault Sacré, Dominique Lafontaine et Marie-Christine Toczek

Le rôle des composantes de la qualité de l'environnement éducatif à la petite enfance sur la préparation scolaire des enfants issus de familles immigrantes

Sabrina Bolduc et Suzanne Manningham

Pensée sociale et adaptation socioscolaire en classe d'éducation préscolaire 5 ans: portrait d'enfants québécois à l'aube de leur parcours scolaire

Sarah Landry, Caroline Bouchard et Pierre Pagé

... et trois autres articles

UDS Université de
Sherbrooke

Comprendre et concevoir des revues systématiques de la littérature en sciences de l'éducation et de la formation

Margault Sacré

Université Clermont Auvergne

Dominique Lafontaine

Université de Liège

Marie-Christine Toczek

Université Clermont Auvergne

Résumé

Les revues systématiques de la littérature proposent une synthèse des connaissances scientifiques sur une question de recherche. La transparence de la méthodologie permet de juger la fiabilité de la revue et des conclusions proposées. L'éducation fondée sur les preuves défend cette méthodologie, mais ignore généralement les études non expérimentales. L'objectif du présent article est de permettre la compréhension et la conception de revues systématiques en sciences de l'éducation et de la formation, incluant différents types de recherche, des études expérimentales aux études à visée compréhensive et exploratoire. Cet article détaille les étapes de la construction d'une revue systématique et propose un script de rédaction pour tout chercheur s'engageant dans l'écriture d'une revue systématique, quel que soit son niveau d'expertise.

Mots-clés

revue systématique, démarche méthodologique, sciences de l'éducation, pratiques éducatives, éducation fondée sur les preuves

Understanding and designing systematic literature reviews in education and training

Abstract

Systematic literature reviews offer a synthesis of scholarly knowledge on a given research subject. The transparency of the methodology is a benchmark of the reliability of the review and the proposed conclusions. Evidence-based education defends this methodology, but generally ignores non-experimental studies. The purpose of this article is to help understand and design systematic reviews in education and training, including different types of research, from experimental studies to comprehensive and exploratory studies. The article details the steps involved in the construction of a systematic review and proposes a writing script for any researcher engaging in writing a systematic review, irrespective of their level of expertise.

Keywords

systematic review, methodological approach, education science, educational practices, evidence-based education

Comprender y diseñar revisiones sistemáticas de literatura en ciencias de la educación y la formación

Resumen

Las revisiones sistemáticas de la literatura proporcionan una síntesis del conocimiento científico sobre una cuestión de investigación. La transparencia de la metodología permite juzgar la fiabilidad de la revisión y de las conclusiones propuestas. La educación basada en evidencias defiende esta metodología, pero generalmente ignora los estudios no experimentales. El objetivo de este artículo es permitir la comprensión y el diseño de revisiones sistemáticas en las ciencias de la educación y la formación, incluyendo diferentes tipos de investigación, desde estudios experimentales hasta estudios integrales y exploratorios. Este artículo detalla los pasos para crear una revisión sistemática y ofrece un guión de redacción para cualquier investigador que participe en la redacción de una revisión sistemática, independientemente de su nivel de experiencia.

Palabras claves

revisión sistemática, enfoque metodológico, ciencias de la educación, prácticas educativas, educación basada en evidencia

1. Introduction

À côté des revues narratives de recherche, dont la fonction principale est de cerner les zones d'ombre dans la façon dont un problème est construit dans un domaine de la littérature, a émergé avec le mouvement de l'éducation basée sur les preuves (EBP) une nouvelle forme de revue de recherche: la revue systématique de la littérature (Saussez et Lessard, 2009). Cette forme de revue est présentée par les tenants de l'EBP comme un moyen neutre et objectif de synthétiser objectivement des résultats de recherche concernant une problématique en suivant un protocole reproductible (Maison, 2010; Oakley, 2002). L'engouement pour cette méthode de synthèse est notamment justifié par la quantité exponentielle des résultats de recherche et la difficulté pour les personnes praticiennes, chercheuses et décisionnaires d'appréhender et de comprendre toutes ces informations (Saussez et Lessard, 2009). Or, une condition de l'EBP réside dans l'existence de synthèses scientifiquement valides et facilement interprétables des recherches sur des programmes éducatifs¹ (Slavin, 2008).

Cette procédure repose généralement sur la compilation et la comparaison d'études expérimentales et plus précisément les essais contrôlés randomisés. S'inspirant de la médecine basée sur les preuves, les chercheurs et chercheuses en sciences de l'éducation et de la formation ont compris que ces études sont précieuses dans le sens où elles permettent de mettre en évidence des relations causales et donc d'attester des effets de programmes ou de pratiques d'enseignement (Slavin, 2002). Dans le champ de l'éducation, ce type d'études n'a pas l'exclusivité et il est tout aussi admis que d'autres types de recherche, telles que les études transversales ou les recherches qualitatives, font partie intégrante du domaine, même s'ils n'apportent pas le même niveau de preuve (Slavin, 2002). Pourtant, ces types de recherches sont généralement omis dans l'EPB (Baye *et al.*, 2016), font beaucoup moins l'objet de revues systématiques (Dixon-Woods *et al.*, 2006)

1 Traduction de: «the existence of scientifically valid and readily interpretable syntheses of research on practical, replicable education programs» [...] (p. 5).

et figurent rarement dans les outils méthodologiques. Il ne s'agit en aucun cas de renier les différences méthodologiques et fonctionnelles des types de recherche des sciences de l'éducation et la formation (Gough, 2004), mais la conception de revues systématiques de la littérature incluant des types de recherche plus variés pourrait venir nourrir les recherches menées dans le cadre de l'EBP. En outre, comme le soulignent Cook et Cook (2016), la variété des types de recherche correspond à la variété des objectifs de recherche.

Fort de ces constats, il apparaît important de disposer d'un écrit de type méthodologique permettant la compréhension, l'utilisation et la conception de revues systématiques de la littérature dans le domaine des sciences de l'éducation et de la formation, englobant à la fois la recherche expérimentale, mais également d'autres types de recherche quantitatifs et qualitatifs. C'est précisément l'intention de cet article. En d'autres termes, l'objectif ici est de décrire les différentes étapes de la construction d'une revue systématique de littérature, en s'appuyant sur les recommandations des collaborations Campbell (Kugley *et al.*, 2017; Page *et al.*, 2021) et Cochrane (Higgins *et al.*, 2020; Higgins et James, 2019), du groupe PRISMA (Moher *et al.*, 2009) et en les adaptant au domaine de recherche de l'éducation et de la formation. Notons qu'à l'heure actuelle nous disposons de guides et d'articles scientifiques présentant cette méthode dans différents domaines, par exemple en médecine, en ergothérapie, en sciences de la santé ou en kinési-physiothérapie (Bussièrès, 2018; Cronin *et al.*, 2008; Gopalakrishnan et Ganeshkumar, 2013; Pollock et Berge, 2018; Regnaud et Remondière, 2018; Zugg *et al.*, 2014). Néanmoins, à notre connaissance aucun article scientifique ne présente cette méthode en sciences de l'éducation et de la formation alors que le paradigme de l'éducation basée sur les preuves est en plein essor dans ce champ au niveau international (Rey, 2006; Slavin, 2008). Finalement, il semble nécessaire de fournir cette procédure pour réaliser une revue systématique de la littérature à un lectorat francophone.

2. Étapes et procédures pour réaliser une revue systématique de la littérature

S'inspirant des définitions conçues, Newman et Gough (2020) définissent les revues systématiques comme des revues de la littérature réalisées par le biais de méthodes explicites. Evans et Benefield (2001) ajoutent que différents éléments clés doivent être explicités: la question de recherche, les méthodes d'identification et de sélection des études incluses à la revue, ainsi que les critères d'inclusion et les éléments choisis pour déterminer la qualité méthodologique des études.

Ainsi, la méthodologie d'une revue systématique de littérature est présentée, d'une part, comme la construction d'une synthèse de travaux de recherche en réponse à une question de recherche et, d'autre part, comme un élément clé de sa qualité grâce à l'explicitation de sa procédure de synthèse, permettant sa réplication (Oakley, 2002). De plus, cette méthodologie comprend la sélection et l'analyse d'études à inclure dans la recherche pour répondre à une question et peut utiliser des méthodes statistiques (méta-analyses) pour synthétiser les résultats (Moher *et al.*, 2009). En somme, une revue systématique doit impliquer «une série de techniques visant à minimiser les biais et les erreurs» (MacDonald, 2000, p. 131).

La construction et la rédaction d'une revue systématique de littérature en sciences de l'éducation et de la formation peut être présentée en cinq étapes: (1) formulation de la question de recherche, (2) définition des critères d'inclusion, (3) identification des études, (4) extraction des données et (5) synthèse des données (Gough, 2007). Dans cet article, nous proposons de présenter successivement chacune de ces étapes. Les étapes présentées sont une traduction et une adaptation des éléments essentiels parus dans différents outils, appuyés d'exemples tirés des sciences de l'éducation et de la formation.

2.1 Formulation de la question de recherche

Cette première étape est fondamentale puisqu'il s'agit de rédiger de façon claire et précise une ou plusieurs questions de recherche qui vont conditionner l'ensemble de la revue systématique (Jahan *et al.*, 2016). Une question trop précise ne pourrait pas conduire les analystes² à une quelconque généralisation et, à l'inverse, une question trop large pourrait entraîner des problèmes d'interprétation (Zaugg *et al.*, 2014). Pour faciliter la formulation de la question, les analystes peuvent s'inspirer de différents cadres de référence, inspirés du modèle PICO (population/problème, intervention, comparaison, objectifs) (Richardson, 1995, cité par Davies, 2011). Davies (2011) recense une panoplie de cadres de référence qui permettent une structuration des éléments à inclure dans la recherche, dans lesquels PICO devient PICOC si l'on veut inclure le contexte de la problématique ou PICOT si l'on veut ajouter un critère de temps. D'autres éléments peuvent être inclus dans les questions: le moment de l'intervention, les professionnels impliqués dans l'intervention, les résultats attendus ou le type d'étude.

Un exemple tiré de Swanson *et al.* (2017) d'adaptation et d'utilisation de ce modèle est proposée dans le tableau 1. Selon ces chercheurs, les effets de l'apprentissage en équipe ont peu été évalués dans l'enseignement supérieur.

2 Nous utilisons dans cet article le terme d'analyste, au sens large, pour faire référence aux personnes concevant la revue systématique.

Tableau 1

	Exemple (Swanson <i>et al.</i> , 2017)
Problème	Peu d'évaluations des effets de l'apprentissage en équipe
Participants	Étudiants et étudiantes de l'enseignement supérieur
Intervention (ou objet de recherche)	L'apprentissage en équipe
Objectif visé	Améliorer les connaissances de contenu

Modèle PICO et exemple.

Ils posent alors les questions de recherche suivante: «Quels sont les effets de l'apprentissage en équipe auprès des étudiants et étudiantes de l'enseignement supérieur sur l'amélioration de leurs connaissances de contenu? Quelles conditions modèrent les effets de l'apprentissage en équipe?», reprenant ainsi les différents éléments du cadre de référence choisi.

Enfin, pour définir une question de recherche, il convient de vérifier qu'elle n'ait pas été précédemment traitée d'un point de vue scientifique. Des bases de données et outils de recherche relatifs aux travaux publiés dans le champ des sciences de l'éducation et de la formation (p. ex., EPPI Centre, Campbell Collaboration, Review of Educational Research, What Works Clearinghouse) sont indispensables pour une telle vérification.

2.2 Définition des critères d'inclusion

Une fois la question de recherche formulée, il est nécessaire de préciser les critères d'inclusion et d'exclusion des études, c'est-à-dire les critères sur la base desquels les études seront incluses dans la synthèse: le *design* des études (la durée, le type d'étude...), des précisions sur la population étudiée (âge, nombre, système éducatif...) ou le langage de publication (Page *et al.*, 2021). Chacun des critères doit être défini de manière explicite.

Par exemple, si les analystes veulent réaliser une revue systématique concernant les effets du statut socioéconomique des étudiants et étudiantes sur leur réussite, ils doivent définir s'ils veulent que les études incluses comprennent les étudiants et étudiantes provenant de l'enseignement postsecondaire non supérieur ou uniquement de l'enseignement supérieur³. Les analystes doivent aussi définir le statut socioéconomique ainsi qu'une échelle de mesure correspondant. Ainsi, si les analystes ont défini le statut socioéconomique sur la base du diplôme des parents, les études ne comportant pas cette mesure seront exclues.

3 Termes en référence à la classification internationale type de l'éducation, CITE 2011 (Institut de statistique de l'UNESCO, 2013).

Les analystes incluent généralement des études sur la base de leurs propres compétences linguistiques, c'est-à-dire que s'ils parlent anglais et français, ils vont exclure toutes les études qui ne sont pas publiées dans ces langues. Cette restriction doit être précisée et indiquée dans les limites d'une revue systématique (Page *et al.*, 2021), dans la mesure où l'inclusion d'études en d'autres langues que les langues anglaise et française permet d'accroître la précision des résultats de la revue systématique (Lefebvre *et al.*, 2019) et donc la fiabilité des conclusions.

Le niveau de rigueur des études à inclure est à adapter en fonction du type des études identifiées. Si on dispose d'articles évaluant des pratiques sous la forme d'études expérimentales ou quasi expérimentales, une méta-analyse peut être réalisée. Par exemple, les méta-analyses publiées par la *Best Evidence Encyclopedia* (BEE, 2021) incluent uniquement des études comparatives. En sciences de l'éducation et de la formation, un grand nombre d'objets de recherche font plutôt l'objet d'étude de cas (qualitatives ou transversales) et, dans ce cas, une revue de littérature est tout aussi envisageable (Fitzpatrick-Lewis *et al.*, 2009). Par exemple, le sujet traité par la revue systématique de Fitzpatrick et Burns (2019) ne comprend aucune étude expérimentale et les auteurs ont ainsi adapté le type d'étude inclus. Ils ont inclus trois types d'études: (1) comparaison d'une école à une autre école similaire, (2) comparaison d'une école à une école voisine, et (3) comparaison d'élèves d'une même école en début et en fin d'année⁴.

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. Cependant, la fiabilité des conclusions des analystes reste dépendante du choix de ces critères. Comme l'ont montré Cheung et Slavin (2016), il existe une relation significative entre les caractéristiques méthodologiques des études incluses dans une revue systématique et les effets rapportés. Plusieurs caractéristiques méthodologiques auraient ainsi tendance à gonfler les résultats et, partant, à influencer les conclusions des analystes. Il s'agit notamment d'être prudent quant aux effets obtenus pour les études ayant moins de 250 sujets, sans répartition aléatoire et avec des variables dépendantes non standardisées. Cheung et Slavin (2016) ont montré que l'exclusion de ces critères peut engendrer des résultats ou des tailles d'effet jusqu'à deux fois plus élevés que lorsque ces critères sont adoptés.

4 Traduction de: «(a) comparing one school to another that is very similar, (b) comparing one school to a nearby school, and (c) comparing students at a single school before versus after a switch to a year-round calendar» (p. 2).

Le choix des critères d'inclusion appartient aux analystes et dépend fortement de la disponibilité des études, mais il reste nécessaire d'expliquer les limites associées à certaines caractéristiques méthodologiques.

2.3 Identification des études

2.3.1 Outils de recherche bibliographiques à consulter

Étant donné qu'aucune base de données bibliographiques ne couvre l'ensemble des publications scientifiques internationales, il apparaît indispensable de s'appuyer sur plusieurs outils de recherche bibliographiques (Zaugg *et al.*, 2014).

Dans le domaine des sciences de l'éducation et de la formation, plusieurs bases de données peuvent être consultées. Spécifiques au domaine, il existe ERIC (Education Resources Information Center), ERA (Educational Research Abstracts Online) ou Érudit, mais une recherche dans d'autres bases de données bibliographiques spécialisées en sciences sociales, humaines et en psychologie, telles que PsycInfo, PsycARTICLES, PsycEXTRA, ISIDORE et Social Science Citation Index, peut parfois s'avérer nécessaire. Des bases de données générales telles que ASP (Academic Search Premier), JSTOR (Journal Storage) et Scopus peuvent également être consultées. De manière à juger les bases de données les mieux adaptées aux critères de recherche, les analystes peuvent vérifier les revues scientifiques qui y sont indexées. Pour améliorer l'exhaustivité de la recherche, les analystes peuvent aussi utiliser des moteurs de recherche plus généraux, comme Google Scholar.

Notons qu'après avoir réalisé une recherche dans les bases de données bibliographiques, il est conseillé de réaliser une recherche inversée (Cited References Search), c'est-à-dire que les analystes vont chercher les études qui citent les études incluses dans leur revue systématique. Des outils de recherche bibliographiques permettent la recherche de citations (Web of Science, Google Scholar).

Comme le rappellent Cheung et Slavin (2016), il est également intéressant de faire une recherche de la littérature grise⁵, c'est-à-dire des rapports non publiés, de manière à réduire un éventuel biais de publication qui serait dû au fait que les effets rapportés dans les rapports non publiés sont généralement

5 «La littérature grise correspond à tout type de document produit par le gouvernement, l'administration, l'enseignement et la recherche, le commerce et l'industrie, en format papier ou numérique, protégé par les droits de propriété intellectuelle, de qualité suffisante pour être collecté et conservé par une bibliothèque ou une archive institutionnelle, et qui n'est pas contrôlé par l'édition commerciale» (Schöpfel, 2012, p. 19).

moins élevés que les effets rapportés dans les études publiées (Cheung et Slavin, 2016; Slavin, 1995). La littérature grise peut recouvrir des actes de conférence, des publications gouvernementales, des thèses ou des rapports, tout aussi intéressants dans une revue systématique (Rosenthal, 1991). OpenGrey est une base de données spécialisée dans la littérature grise. À cet effet, les rapports de l'Open DOAR (Directory of Open Access Repositories) et du DOAJ (Directory of Open Access Journals) peuvent également être consultés, ainsi que les bases de données qui archivent les thèses, telles que Dissertation Abstracts International, ProQuest Dissertations & Theses Global, Thèses.fr, ou encore les archives ouvertes HAL.

Enfin, la lecture des études incluses et de revues de littérature préexistantes peut amener à l'identification de documents supplémentaires passés sous le radar et pertinents pour la revue systématique (Cronin *et al.*, 2008).

L'interrogation de ces différents outils peut nécessiter l'aide d'une personne spécialisée en documentation afin de s'assurer de la pertinence des recherches. Ceci est d'autant plus important que chaque base et outil de recherche bibliographique nécessite une méthodologie de recherche particulière.

2.3.2 Planification des requêtes et identification des études

Même s'il n'existe pas de stratégie idéale, le but de la planification des requêtes est de diminuer au maximum les biais de sélection et de réaliser une recherche la plus exhaustive possible afin de rendre compte d'un maximum d'études, tout en tenant compte des critères d'inclusion.

La stratégie de recherche s'appuie sur une liste de mots-clés et de synonymes qui sont généralement inventoriés à partir des articles de référence. Cette liste de mots-clés émerge de la question de recherche et des éléments précisés dans le modèle PICO ou l'un de ses dérivés (Kugley *et al.*, 2017). En identifiant chaque élément du modèle et ses synonymes, une équation de recherche pourra être élaborée à l'aide des opérateurs booléens (AND, OR, NOT).

Par exemple, Vlachopoulos et Makri (2017) ont examiné les effets de l'apprentissage par le jeu vidéo dans l'enseignement supérieur à travers une revue systématique. Les mots-clés utilisés dans leur équation de recherche sont présentés dans le tableau 2.

Tableau 2

Éléments inclus dans l'équation	Mots-clés utilisés
Public ciblé	<i>higher education</i>
Diffusion de la pratique	<i>computer-based, web-based, digital, virtual, online, technology</i>
Objet de recherche	<i>educational games, business simulations, role-playing simulations, game-based learning, video games, serious games</i>
Objectif visé	<i>learning objectives, learning goals, learning objectives, effects</i>

Mots-clés de l'équation de recherche conçue par Vlachopoulos et Makri (2017).

Quatre éléments constituent cette équation: l'objet de recherche et sa méthode de diffusion, le public ciblé et l'objectif. Sur la base de ces éléments et des mots-clés identifiés, l'équation de recherche construite est la suivante:

higher education AND (educational games OR business simulations OR role-playing simulations OR game-based learning OR video games OR serious games) AND (learning objectives OR learning goals OR learning objectives OR effects) AND (computer-based OR web-based OR digital OR virtual OR online OR technology)

D'une base de données à l'autre, les équations de recherche peuvent varier car elles n'utilisent pas toutes les mêmes terminologies. Les analystes doivent donc vérifier que le vocabulaire qu'ils utilisent correspond au vocabulaire de la base de données. De plus, certaines bases de données comprennent un *thesaurus*, c'est-à-dire, un répertoire structuré de mots-clés, permettant aux analystes de voir des termes connexes à leurs mots-clés et, dès lors, de compléter leur équation de recherche (Kugley *et al.*, 2017).

L'élaboration d'une équation de recherche se fait de manière itérative, en fonction du bruit et du silence documentaire engendrés par la stratégie de recherche. Le silence documentaire, «ensemble des documents pertinents non affichés lors d'une recherche documentaire», est généralement causé par l'absence d'un mot-clé ou d'un synonyme inhérent à la recherche ou par un terme trop spécifique au domaine ou au contexte. À l'inverse, le bruit documentaire, «ensemble des documents non pertinents affichés suite à une recherche documentaire», est causé par des mots-clés polysémiques ou trop généraux (Hensens, 2002, s. p.).

Dans certaines bases de données, il est possible d'utiliser des filtres, c'est-à-dire des stratégies conçues pour retrouver des documents sur la

base, par exemple, de la méthodologie. Cependant, les bases de données en sciences sociales ont tendance à ne pas être aussi systématiquement indexées que les bases de données en sciences de la santé, donc l'utilisation des filtres n'est pas toujours recommandée (Kugley *et al.*, 2017). Des limites peuvent également être appliquées, pour restreindre les résultats en fonction du type, de l'année ou de la langue de publication.

Lorsque les analystes s'intéressent aux pratiques éducatives, les variations terminologiques dans les systèmes éducatifs à travers le monde exigent une grande vigilance. En effet, d'un pays à l'autre, un même terme ne désignent pas le même niveau de formation. Par exemple, en France, le terme «collège» concerne les élèves de 11 à 15 ans, alors que le terme *college* aux États-Unis correspond aux étudiants et étudiantes universitaires. Lors de l'étape de la planification des requêtes, il est conseillé d'identifier les termes utilisés dans les différents pays où la pratique est étudiée. À ce propos, la classification internationale type de l'éducation (Institut de statistique de l'UNESCO, 2013) représente un outil pertinent.

2.3.3 Criblage et sélection des études

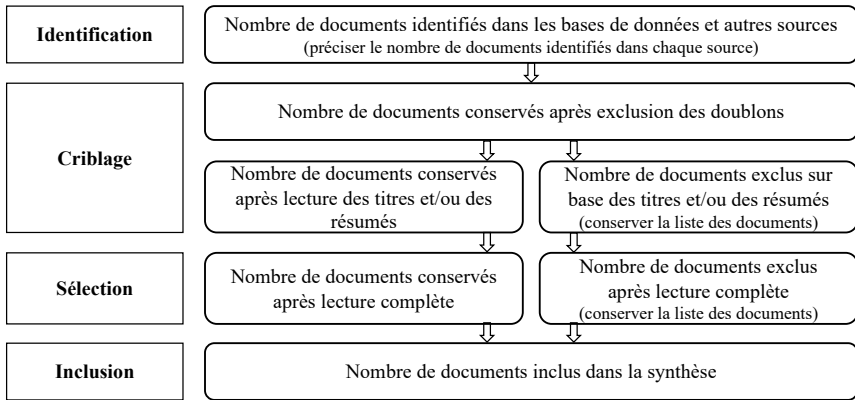
Une fois la recherche d'articles réalisée, l'analyste peut importer les références identifiées dans un logiciel bibliographique (p. ex., EndNote, Zotero ou Mendeley). Après avoir supprimé les doublons, il convient de procéder à un criblage qui repose sur une analyse sémantique des titres et résumés. Ainsi, l'analyste élimine les articles qui seraient trop éloignés de la question de recherche.

Ensuite, les articles les plus pertinents par rapport aux questions de recherche sont téléchargés et lus en intégralité pour déterminer leur pertinence au regard des critères définis précédemment. S'ils sont jugés pertinents, ils seront comptabilisés comme «articles éligibles».

Lors de la sélection des travaux de recherche, il est important de lister les études exclues ainsi que les arguments qui président à cette décision de non-éligibilité. Pour faciliter l'organisation de ce travail de criblage et de sélection, l'utilisation d'un outil d'aide spécifique à la revue, tel que l'application libre Rayyan (Ouzzani *et al.*, 2016), est conseillée.

L'ensemble de ce processus est souvent représenté par un diagramme de flux (Moher *et al.*, 2009). À titre d'exemple, nous proposons en figure 1 un modèle de diagramme de flux.

Figure 1. Exemple d'un diagramme de flux représentant les différentes phases d'une revue systématique adapté de Moher *et al.* (2009).



Les analystes doivent préciser les bases de données bibliographiques et tous les autres outils de recherche utilisés, ainsi que les équations de recherche. Les dates auxquelles chaque source a été utilisée sont également indiquées pour permettre aux analystes qui souhaitent mettre à jour la revue systématique d'avoir une date de référence à partir de laquelle les documents nouvellement publiés ne sont pas inclus.

Bien que cela ne soit pas toujours possible d'un point de vue pratique et financier, le criblage et la sélection des études peuvent être réalisés par deux juges de façon indépendante pour en assurer la fiabilité (Bernard *et al.*, 2014). Cette précaution méthodologique doit s'accompagner d'une procédure de conciliation pour résoudre les éventuels désaccords.

3. Extraction des données

3.1 Repérage des données pertinentes pour la revue systématique

À ce stade, toutes les études incluses dans la revue systématique sont identifiées et les analystes doivent procéder à l'extraction des données, c'est-à-dire les éléments caractérisant les études. Les données nécessaires à l'analyse des études peuvent être recueillies à l'aide d'un canevas. Pour ce faire, des documents types existent, comme le Cochrane data collection form for intervention review (The Cochrane Collaboration, 2020) ou le PRISMA Statement (Moher *et al.*, 2009). Ces formulaires peuvent être utilisés en sciences de l'éducation et de la formation en étant adaptés et testés par les analystes selon les objectifs de la revue systématique.

Les éléments suivants peuvent constituer le canevas:

- Objets de recherche, définition et objectifs;
- Composantes des objets de recherche: contenu, données géographiques, dates;
- Type d'étude: étude prospective ou rétrospective;
- Durée de l'étude;
- Échantillon;
- Taille d'échantillon;
- Niveau scolaire des participants;
- Caractéristiques des participants: âge, sexe, statut social, origine, langue maternelle, appartenance à une minorité;
- Concepts ou théories du cadre de référence;
- Méthode de recrutement des participants dans l'étude;
- Variables indépendantes et échelles de mesure;
- Conception de l'étude: *design* de l'étude, méthode d'analyse des données, réplicabilité;
- Résultats;
- Risques de biais.

Ces éléments doivent être définis avec une grande précision car la qualité méthodologique des analyses réalisées et des conclusions tirées est étroitement dépendante de l'utilisation d'un tel canevas. Des logiciels de gestion de données (p. ex., Microsoft Excel) ou de codage de textes (p. ex., QDAMiner ou NVivo) permettent l'extraction et la transcription d'un grand nombre de données ainsi que la conception d'un tableau à double entrée pour y accéder facilement lors de leur synthèse.

3.1.2 Évaluation de la qualité méthodologie des études

L'évaluation de la qualité méthodologique des études représente la pierre angulaire de la revue systématique car l'ensemble de ses conclusions en dépend. Aussi, cette évaluation sera-t-elle réalisée en filigrane à l'ensemble des étapes caractérisant l'analyse des résultats, car si des problèmes de conception ou de réalisation n'étaient pas détectés ils produiraient de fausses interprétations et biaiseraient les conclusions (Cheung et Slavin, 2016). Le tableau 3, inspiré de Tripney *et al.* (2015, p. 94) et adapté à différents types de recherche, propose une liste des risques de biais à identifier avant et pendant l'analyse des résultats.

Tableau 3

Biais	Définition du biais	Questions
Biais de sélection	Les biais conduisant à ce que les personnes participantes à l'étude ne constituent pas un groupe représentatif de la population étudiée (Goldberg, 2011; Norris, 1997).	La conception de l'étude ou l'analyse des résultats tient-elle compte d'un possible biais de sélection?
Biais d'attrition	Les biais dus à la perte de personnes participantes durant une étude.	Les données manquantes et/ou incomplètes (non-réponse, abandon ou exclusion de personnes participantes) sont-elles traitées?
Biais de performance	Les biais rencontrés lorsque les personnes participantes et le personnel chargé des pratiques ne sont pas traités de manière équivalente (Higgins <i>et al.</i> , 2011; Martineau, 2005).	Sauf dans le cas d'une étude comparative, les personnes participantes ont-elles été traitées sur un pied d'égalité à tous les égards? Dans le cas d'une étude comparative, la connaissance de l'attribution aux groupes témoin et expérimental a-t-elle été évitée?
Biais de détection	Les biais rencontrés lorsque toutes les étapes du processus ne sont pas mises en place de manière consistante (Higgins <i>et al.</i> , 2011; Norris, 1997)	Les résultats ont-ils été évalués à l'aide de mesures valables et fiables, et mis en œuvre de manière consistante pour toutes les personnes participantes? Dans le cas d'une étude comparative, les évaluateurs des résultats ont-ils été aveugles au statut d'intervention des personnes participantes?
Biais de rapport	Les biais dus à la sélection des résultats rapportés (Chevalier, 2014; Martineau, 2005).	Tous les résultats des mesures prédéfinies sont-ils communiqués? Les analyses supplémentaires sont-elles justifiées?

Biais à identifier et questions à poser pour les identifier.

Il s'agit également d'identifier si l'influence des chercheurs sur l'étude a été prise en compte, si les personnes participantes sont représentées de manière adéquate, si la méthodologie choisie est appropriée pour répondre

aux objectifs de l'étude et si cette méthodologie est respectée dans la collecte, la présentation et les analyses des données (Lockwood *et al.*, 2015).

Il est important de détecter les biais des études incluses et d'agir en conséquence, c'est-à-dire, en nuancant leurs conclusions, voire en les éliminant de la procédure.

4. Synthèse des données

La synthèse est réalisée à partir des données extraites de chaque étude. Les analystes sont conduits à décrire les objets de recherche étudiés, à synthétiser les données extraites et, éventuellement à les catégoriser. Dans une revue systématique, les données peuvent être synthétisées de manière descriptive, semi-quantitative ou quantitative. Lorsque les analystes souhaitent inclure les études de cas, les études exploratoires et les études descriptives dans leur revue systématique, ils peuvent opter pour une synthèse descriptive ou semi-quantitative (Dixon-Woods *et al.*, 2006; Zaugg *et al.*, 2014). L'analyse quantitative est envisageable lorsque l'on travaille à partir d'un corpus d'études relativement homogènes et lorsque l'on souhaite faire une méta-analyse, incluant alors des études comparatives.

4.1 Synthèse descriptive

Dans un premier temps, la synthèse descriptive consiste en un exposé des caractéristiques des études présentant la méthodologie de chaque étude, l'objet de la recherche et ses résultats. Dans un second temps, il s'agit d'associer ou de catégoriser les études de manière logique dans le but d'orienter les conclusions de la revue. Dès lors, si l'objectif de la revue est de rendre compte de l'efficacité d'une pratique éducative en comparaison avec d'autres, alors une catégorisation des études en fonction du type de pratique ou des résultats qui révèlent l'efficacité de plusieurs types de pratique sera privilégiée. En revanche, si l'objectif de la revue est de déterminer quel est le contexte le plus favorable à tel ou tel comportement, alors c'est une catégorisation par contexte ou par population qui permettra de répondre à l'objectif fixé. Et enfin, si l'objectif de la revue est de révéler les études les plus probantes dans un domaine, alors la catégorisation sera réalisée en fonction de critères méthodologiques convoqués dans les études retenues. Dans un troisième temps, une synthèse des résultats est réalisée à partir du travail précédent.

Pour étudier les attitudes des personnes enseignant l'éducation physique envers l'inclusion, Tant et Watelain (2014) ont opté pour une

analyse qualitative. Leur procédure est réalisée selon l'analyse inductive (Blais et Martineau, 2006; Thomas, 2006), c'est-à-dire, qu'après une lecture approfondie des études, l'identification de catégories est réalisée par les auteurs pour revenir ensuite aux données brutes de chaque étude. Les catégories émergentes incluent notamment l'expérience des personnes enseignantes, leur sentiment de compétence vis-à-vis de l'inclusion, leur formation ou le type de handicap. Les données sont alors détaillées dans chaque catégorie et les liens entre les différentes études sont explicités.

Tondeur *et al.* (2017) ont conduit une revue systématique de la littérature incluant spécifiquement des études qualitatives, avec pour objectif de mettre en évidence les liens entre les croyances des personnes enseignantes et leur utilisation des technologies. Ils proposent alors une catégorisation des données, incluant, entre autres, les technologies perçues comme catalyseur de changement, les freins liés à l'utilisation des technologies, la résistance au changement, ou encore l'influence de l'environnement éducationnel. Leur revue systématique permet alors de détailler plusieurs thématiques liées à leur problématique. D'après ces auteurs, les croyances des personnes enseignantes peuvent constituer une barrière à l'intégration des technologies, alors que l'intégration des technologies aurait le potentiel de transformer ces croyances, notamment à travers une formation continue ciblée et des politiques de l'école.

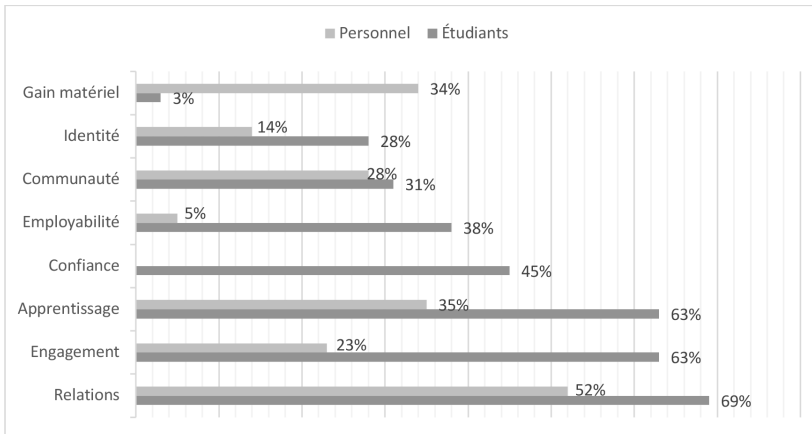
De telles synthèses représentent une réelle plus-value scientifique puisqu'elles offrent un travail de synthèse méthodique (Bondas et Hall, 2007; Dixon-Woods *et al.*, 2006). Elles permettent, entre autres, «d'identifier des zones d'ombre dans la recherche» (Saussez et Lessard, 2009, p. 113) et de mettre alors en évidence de nouvelles perspectives de recherches. Pourtant, les analystes doivent rester conscients des biais possibles de ce type de synthèse, qui peuvent être incomplètes, trop théorisées, pas assez synthétisées ou présenter une catégorisation des études sans fondement explicite (Bondas et Hall, 2007). Par ailleurs, selon Thomas et Harden (2008), les analystes tombent parfois dans le piège de décontextualiser des données à l'extrême et, partant, de comparer des données non comparables. Cela peut notamment être dû au fait que, dans les synthèses descriptives, les analystes ne retournent pas aux données ou constats primaires des études, mais s'appuient sur les interprétations et les conclusions tirées dans les études incluses (Estabrooks *et al.*, 1994). Les analystes optant pour ce type de synthèse doivent être conscients des limites possibles pour les minimiser et en rendre compte dans leur revue systématique.

4.1 Synthèse semi-quantitative

Les analystes peuvent opter pour une synthèse semi-quantitative, c'est-à-dire que des données quantitatives peuvent être présentées, sans pour autant révéler des tailles d'effet. Dans ce cas de figure, les analystes vont, par exemple, indiquer si la pratique provoque un effet positif, négatif ou s'il ne provoque pas d'effet sur les mesures ciblées.

Un exemple de ce type de synthèse est présenté dans la revue systématique de Matthews *et al.* (2019). Les auteurs ont construit un graphique (figure 2) avec, en ordonnée, les mesures ciblées et, en abscisse, le pourcentage d'études montrant un effet positif de la pratique pour les élèves et pour le personnel.

Figure 2. Études faisant état d'effets positifs sur la collaboration entre les étudiants et étudiantes et le personnel, en pourcentage, d'après Matthews *et al.* (2019, p. 6).



Un autre exemple, tiré de la littérature francophone, décrit en pourcentages les effets de l'usage de portfolios sur l'apprentissage des étudiants et des étudiantes (tableau 4). Buckley et ses collègues (2012) classent les effets signalés selon qu'ils sont positifs, négatifs ou inexistantes. Ces effets sont mis en parallèle avec l'objectif d'apprentissage.

Tout comme les synthèses descriptives, ce type de synthèse ne montre pas les effets réels des objets de recherche étudiés, mais offre plutôt un aperçu des études disponibles et de leurs résultats. La réalisation d'un travail de ce type est donc importante mais n'est pas exempte de biais.

Tableau 4. Les effets de l'usage de portfolios sur les apprentissages d'après Buckley *et al.* (2012, p. 130).

Objectifs d'apprentissage	Exemples	Études signalant des effets (%)		
		Positifs	Aucun	Négatifs
Connaissance et compréhension	Objectifs des cours atteints	29	1	
	Intégration de la théorie à la pratique	12	4	
Habilités	Réflexivité/connaissance de soi	64	4	
	Gestion des patients	4	1	
	Prise de décision	9		
Attitudes et comportements	Confiance en soi	12		1
	Responsabilité	7		

Lecture du tableau: 29 % des études incluses dans la revue systématique rapportent des effets positifs sur l'atteinte des objectifs de cours.

Tout comme les synthèses descriptives, ce type de synthèse ne montre pas les effets réels des objets de recherche étudiés, mais offre plutôt un aperçu des études disponibles et de leurs résultats. La réalisation d'un travail de ce type est donc importante mais n'est pas exempte de biais.

4.2 Synthèse quantitative ou méta-analyse

Pour estimer les effets des pratiques, des interventions ou des programmes de recherche, les analystes peuvent réaliser une méta-analyse si les données recueillies leur permettent de calculer les tailles d'effet. D'après Lipsey et Wilson (1999), tout indice standardisé peut être une taille d'effet pour autant qu'il réponde aux critères suivants:

- L'indice est comparable d'une étude à l'autre;
- L'indice représente l'ampleur et la direction de la relation d'intérêt;
- L'indice est indépendant de la taille de l'échantillon.

Dès lors, une différence de moyennes standardisée (Cheung et Slavin, 2012; Pellegrini *et al.*, 2021), un coefficient de corrélation (Hjetland *et al.*, 2017) ou un risque relatif rapproché (*odds ratio*) (p. ex., Fitzpatrick et Burns, 2019) peut être utilisé comme taille d'effet dans une méta-analyse.

Lorsque le calcul de la taille d'effet est possible, il le sera pour chaque pratique et permettra de classer les pratiques incluses. En effet, plus la taille d'effet est élevée, plus l'effet de la pratique sur la variable étudiée est important. Cependant, définition d'une taille d'effet importante ne fait pas consensus dans la communauté scientifique car la qualité du *design* de l'étude peut fortement influencer les tailles d'effet. La BEE (2021) donne l'exemple suivant: une intervention de grande ampleur avec affectation aléatoire au groupe expérimental et au groupe témoin avec une taille d'effet de .20 est plus intéressante qu'une petite intervention avec des échantillons appariés et une taille d'effet de .40⁶. Au-delà de cet argument méthodologique, Hill *et al.* (2008) soutiennent qu'il est nécessaire d'interpréter les effets des pratiques de manière relative, en les comparant à des pratiques similaires, mais également en fonction du contexte et des écarts liés, par exemple, à l'appartenance à une minorité. Pour ces raisons et pour permettre de nuancer les résultats au regard de la fiabilité des études, les éléments méthodologiques de chaque étude doivent être explicitement documentés.

Les résultats peuvent être présentés comme dans le rapport de Slavin *et al.* (2009), proposés dans le tableau 5. Dans ce tableau, différents éléments sont précisés au regard de la taille d'effet et les interventions sont catégorisées: apprentissage coopératif, réforme globale de l'école. De plus, cette méta-analyse documente les mesures dépendantes utilisées dans chaque étude.

Tableau 5. Informations descriptives et tailles d'effet d'études incluses dans la méta-analyse de Slavin *et al.* (2009, p. 106, 109).

Étude	Méthode	Durée	N	Année scolaire	Caractéristiques de l'échantillon	Taille d'effet
Apprentissage coopératif						
Slavin et Karweit (1984)	Étude expérimentale	1 an	588 élèves de 44 classes dans 26 écoles	Secondaire	Élèves faibles à Philadelphie, faible milieu socio-économique	+ .21
Mevarech <i>et al.</i> (2006)	Étude quasi expérimentale	1 semestre	100 élèves dans 4 classes	Secondaire	Élèves israéliens	+ .21
Réforme globale de l'école						
Balfanz <i>et al.</i> (2004)	Étude comparative	1 an	373 élèves dans 6 écoles	Secondaire	Centre-ville de Baltimore	+ .18

6 Traduction de: «a large experiment with random assignment to treatments that obtained an effect size of +0.20 is more important than a small, matched experiment with an effect size of +0.40» (s. p.).

Tableau 6

1. Élaboration de la question de recherche (modèle PICO) et définitions des éléments	
2. Méthodologie	
	Protocole de recherche
	Critères d'inclusion des études
	Outils de recherche consultés
	Stratégies de recherche pour chaque outil de recherche
	Diagramme de flux
	Identification des risques de biais de publication
3. Résultats descriptifs	
	Caractéristiques des études incluses
	Identification des risques de biais dus à la méthodologie des études
	Description des objets de recherche
4. Synthèse des résultats: quantitative ou descriptive	
5. Conclusions	
	Résumé des résultats
	Réponse à la question de recherche au regard des résultats
6. Limites et implications de la revue systématique de littérature	
7. Annexe: documents exclus et raisons de leur exclusion (optionnel)	

Script de rédaction d'une revue systématique de la littérature en sciences de l'éducation et de la formation.

Dans cette perspective, les revues systématiques de la Campbell Collaboration s'appuient en grande majorité sur des études permettant de calculer la taille d'effet. En conséquence, les auteurs qui collaborent avec ce consortium excluent généralement les études pour lesquelles ils ne peuvent pas calculer la taille d'effet et il n'est pas rare de voir des revues systématiques appuyant leurs conclusions sur six études alors qu'elles portaient d'un corpus d'une trentaine d'études (p. ex., Bøg *et al.*, 2018).

5. Script de rédaction d'une revue systématique de littérature

La qualité d'une revue systématique dépend étroitement de sa rigueur méthodologique à toutes les étapes de sa construction et la rédaction du texte n'échappe pas à cette règle. À ce sujet, il existe des guides permettant

de suivre un script rédactionnel, tels que le PRISMA Statement (Moher *et al.*, 2009) et le Methodological Expectations of Cochrane Intervention reviews (Higgins *et al.*, 2020). Largement inspiré de ces guides et adapté au domaine des sciences de l'éducation et de la formation, le tableau 6 présente une liste des éléments nécessaires à la publication d'une revue systématique. La présence de ces éléments rend possible une éventuelle réplique ou mise à jour de la revue systématique et guide les personnes praticiennes et chercheuses dans leur lecture.

Pour la valorisation scientifique de la revue systématique, les analystes seront conduits à prendre en compte des recommandations spécifiques à chaque support de publication et donc, possiblement, à infléchir certains éléments du script proposé ici.

6. Conclusion

Comme dans de très nombreux domaines scientifiques, les sciences de l'éducation et de la formation constatent un rythme effréné de sa production. Réaliser une synthèse des travaux existants sur une question précise représente ainsi un réel défi scientifique et la revue systématique s'avère une méthode susceptible de le relever. De plus, la revue systématique représente un texte à même d'intéresser de nombreuses personnes praticiennes puisque ses conclusions pourront également être investies d'un point de vue professionnel (Saussez et Lessard, 2009).

La construction et la publication d'une revue systématique de littérature s'inscrit dans l'approche de l'éducation basée sur les preuves, qui a pour objectif d'«améliorer la qualité scientifique de la recherche en éducation [et de] privilégier des méthodologies répondant à cet objectif notamment les démarches expérimentales (ou quasi expérimentales) ainsi que les «revues systématiques de recherches»» (Rey, 2006, p. 8). En général, les guides méthodologiques offrent peu d'importance aux types de recherche s'éloignant des recherches expérimentales. Pourtant, en sciences de l'éducation et de la formation, tout comme dans d'autres domaines, il existe des types de recherches variées, tout aussi essentielles pour comprendre les processus des contextes éducatifs, qui méritent d'être associés à la méthode de revue systématique. Par ailleurs, toutes les questions de recherche ne trouvent pas de réponse dans les études expérimentales. En sciences de la santé et en médecine, il existe d'ailleurs un mouvement préconisant les revues systématiques de la littérature incluant différents types d'études, aussi bien quantitatives que qualitatives (Caracelli et Cooksy, 2013; Chalmers, 2005; Dixon-Woods *et al.*, 2005; Gough, 2015).

Il s'agit donc de proposer aux analystes des sciences de l'éducation et de la formation une méthode leur permettant d'inclure différents types d'études, tant qu'elles restent pertinentes pour répondre à leurs objectifs.

Cette proposition n'est pas sans rappeler que les limites des revues systématiques de littérature sont liées aux choix méthodologiques des analystes et donc aux études incluses dans la synthèse. Il convient donc de rester vigilants quant aux conclusions proposées et aux implications éducatives liées à ces conclusions puisque selon la sévérité des critères dépendra la valeur des preuves apportées. Pour contrer ces limites, Slavin (1995) conseille de réaliser des *best-evidence syntheses* plutôt que d'autres types de revues systématiques mais, ce n'est pas toujours envisageable, ni la voie choisie par tous les analystes puisque les études expérimentales ne sont pas toujours adaptées aux contextes éducatifs. Les conclusions d'une revue systématique doivent donc toujours être tirées avec précaution et en tenant compte des biais possibles induits par les critères adoptés.

Enfin, il nous semble nécessaire de rappeler que la méthode de revue systématique est appropriée lorsque l'on engage la réalisation d'une revue exhaustive de la littérature, mais elle n'a pas le monopole de la synthèse. En effet, selon les objectifs des analystes, mais également selon le temps qui leur est accordé, d'autres types de revues peuvent être conçues. À ce propos, nous renvoyons les analystes vers la publication de Grant et Booth (2009) qui propose une analyse de 14 types de revues.

Références

- Baye, A., Bluge, V. et Lafontaine, D. (2016). *L'éducation fondée sur des preuves. Rapport du groupe d'expertise mandaté dans le cadre du Pacte pour un enseignement d'excellence*. Liège: Université de Liège. Repéré à https://orbi.uliege.be/bitstream/2268/252129/1/Baye_Bluge_Education%20fondée%20sur%20des%20preuves_1_10_16.pdf
- Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M. et Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, 26(1), 87-122. doi: <https://doi.org/10.1007/s12528-013-9077-3>
- Best Evidence Encyclopedia (2021, 6 février). About the Best Evidence Encyclopedia (BEE). *Best Evidence Encyclopedia*. <https://bestevidence.org/about/>
- Blais, M. et Martineau, S. (2006). L'analyse inductive générale: Description d'une démarche visant à donner un sens à des données brutes. *Recherches qualitatives*, 26(2), 1-18.
- Bøg, M., Filges, T. et Jørgensen, A. M. K. (2018). Deployment of personnel to military operations: Impact on mental health and social functioning. *Campbell Systematic Reviews*, 14(1), 1-127. doi: <https://doi.org/10.4073/csr.2018.6>

- Bondas, T. et Hall, E. O. C. (2007). A decade of metasynthesis research in health sciences: A meta-method study. *International Journal of Qualitative Studies on Health and Well-Being*, 2(2), 101-113. doi: <https://doi.org/10.1080/17482620701251684>
- Buckley, S., Coleman, J., Davison, I., Khan, K. S., Zamora, J., Malick, S., Morley, D., Pollard, D., Ashcroft, T., Popovic, C. et Sayers, J. (2012). Les effets éducatifs des portfolios sur l'apprentissage des étudiants pendant le cursus prégradué: une revue systématique de la collaboration Best Evidence Medical Education (BEME). Guide BEME N° 11. *Pédagogie médicale*, 13(2), 115-145. doi: <https://doi.org/10.1051/pmed/2012013>
- Bussièrès, E.-L. (2018). La revue systématique d'études quantitatives. *Revue francophone de recherche en ergothérapie*, 4(1), 79-83. doi: <https://doi.org/10.13096/rfre.v4n1.107>
- Caracelli, V. J. et Cooksy, L. J. (2013). Incorporating qualitative evidence in systematic reviews: Strategies and challenges. *New Directions for Evaluation*, 2013(138), 97-108. doi: <https://doi.org/10.1002/ev.20061>
- Chalmers, I. (2005). If evidence-informed policy works in practice, does it matter if it doesn't work in theory? *Evidence & Policy. A Journal of Research, Debate and Practice*, 1(2), 227-242. doi: <https://doi.org/10.1332/1744264053730806>
- Cheung, A. C. K. et Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review*, 7(3), 198-215. doi: <https://doi.org/10.1016/j.edurev.2012.05.002>
- Cheung, A. C. K. et Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283-292. doi: <https://doi.org/10.3102/0013189X16656615>
- Chevalier, P. (2014). *Biais de sélectivité lors du choix et du rapport des résultats*. Minerva Website. Repéré à <http://www.minerva-ebm.be/FR/Article/22>
- Cook, B. et Cook, L. (2016). Research designs and special education research. Different designs address different questions. *Learning Disabilities Research and Practice*, 31. doi: <https://doi.org/10.1111/ldrp.12110>
- Cronin, P., Ryan, F. et Coughlan, M. (2008). Undertaking a literature review: A step-by-step approach. *British Journal of Nursing (Mark Allen Publishing)*, 17(1), 38-43. doi: <https://doi.org/10.12968/bjon.2008.17.1.28059>
- Davies, K. S. (2011). Formulating the evidence based practice question: A review of the frameworks. *Evidence Based Library and Information Practice*, 6(2), 75-80. doi: <https://doi.org/10.18438/B8WS5N>
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B. et Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy*, 10(1), 45-53.
- Dixon-Woods, M., Bonas, S., Booth, A., Jones, D. R., Miller, T., Sutton, A., Shaw, R., Smith, J. et Young, B. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6.

- Estabrooks, C. A., Field, P. A. et Morse, J. M. (1994). Aggregating qualitative findings: An approach to theory development. *Qualitative Health Research*, 4(4), 503-511. doi: <https://doi.org/10.1177/104973239400400410>
- Evans, J. et Benefield, P. (2001). Systematic reviews of educational research: Does the medical model fit? *British Educational Research Journal*, 27(5), 527-541.
- Fitzpatrick, D. et Burns, J. (2019). Single-track year-round education for improving academic achievement in U.S. K-12 schools: Results of a meta-analysis. *Campbell Systematic Reviews*, 15(3), e1053. doi: <https://doi.org/10.1002/cl2.1053>
- Fitzpatrick-Lewis, D., Ciliska, D. et Thomas, H. (2009). *Méthodes pour faire la synthèse d'études sans groupes témoins*. Centre de collaboration nationale des méthodes et outils.
- Goldberg, M. (2011). *L'épidémiologie sans peine*. Québec, QC: Presses de l'Université du Québec.
- Gopalakrishnan, S. et Ganeshkumar, P. (2013). Systematic reviews and meta-analysis: Understanding the best evidence in primary healthcare. *Journal of Family Medicine and Primary Care*, 2(1), 9-14. doi: <https://doi.org/10.4103/2249-4863.109934>
- Gough, D. (2004). Systematic research synthesis to inform the development of policy and practice in education. Dans G. Thomas et R. Pring (dir.), *Evidence-based Practice* (p. 44-62). Buckingham, MK: Open University Press.
- Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213-228. doi: <https://doi.org/10.1080/02671520701296189>
- Gough, D. (2015). Qualitative and mixed methods in systematic reviews. *Systematic Reviews*, 4(1), 181. doi: <https://doi.org/10.1186/s13643-015-0151-y>
- Grant, M. J. et Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91-108. doi: <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Hensens, H. (2002). *Bruit et silence*. Repéré à <https://www.mpl.ird.fr/documentation/indexation/bruit.htm>
- Higgins, J., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L. et Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343. doi: <https://doi.org/10.1136/bmj.d5928>
- Higgins, J. et James, T. (dir.). (2019). *Cochrane handbook for systematic reviews of interventions* (2^e éd.). Oxford: Wiley-Blackwell.
- Higgins, J., Lasserson, T., Chandler, J., Tovey, D., Thomas, J., Flemyng, E. et Churchill, R. (2020). *Methodological expectations of Cochrane intervention reviews. Standards for the conduct and reporting of new Cochrane Intervention Reviews, reporting of protocols and the planning, conduct and reporting of updates*. Cochrane.
- Hill, C. J., Bloom, H. S., Black, A. R. et Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172-177.

- Hjetland, H. N., Brinchmann, E. I., Scherer, R. et Melby-Lervåg, M. (2017). Preschool predictors of later reading comprehension ability: A systematic review. *Campbell Systematic Reviews*, 13(1), 1-155. doi: <https://doi.org/10.4073/csr.201714>
- Institut de statistique de l'UNESCO (2013). *Classification internationale type de l'éducation: CITE 2011*.
- Jahan, N., Naveed, S., Zeshan, M. et Tahir, M. A. (2016). How to conduct a systematic review: A narrative literature review. *Cureus*, 8(11). doi: <https://doi.org/10.7759/cureus.864>
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K. et Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, 13(1), 1-73. doi: <https://doi.org/10.4073/cmg.2016.1>
- Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M.-I., Noel-Storr, A., Rader, T., Shokraneh, F., Thomas, J. et Wieland, L. S. (2019). Searching for and selecting studies. In *Cochrane Handbook for Systematic Reviews of Interventions* (p. 67-107). Hoboken, NJ: John Wiley & Sons. doi: <https://doi.org/10.1002/978119536604.ch4>
- Lipsey, M. W. et Wilson, D. B. (1999). *Practical Meta-Analysis* (Sage). Newbury Park, CA: SAGE Publishing.
- Lockwood, C., Munn, Z. et Porritt, K. (2015). Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, 13(3), 179-187. doi: <https://doi.org/10.1097/xeb.000000000000062>
- Macdonald, G. (2000). Social care: Rhetoric and reality. Dans Davies, H. T. O., Nutley, S. M. et Smith, P. C. (dir.), *What Works? Evidence-based policy and practice in public services*. Oxford: Policy Press.
- Maison, P. (2010). La méta-analyse sur données résumées. *Recherche en soins infirmiers*, 101(2), 18-24.
- Martineau, S. (2005). L'instrumentation dans la collecte des données. *Recherches qualitatives, hors-série*, 2, 5-17.
- Matthews, K. E., Mercer-Mapstone, L., Dvorakova, S. L., Acai, A., Cook-Sather, A., Felten, P., Healey, M., Healey, R. L. et Marquis, E. (2019). Enhancing outcomes and reducing inhibitors to the engagement of students and staff in learning and teaching partnerships: Implications for academic development. *International Journal for Academic Development*, 24(3), 246-259. doi: <https://doi.org/10.1080/1360144X.2018.1545233>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. et The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7). doi: <https://doi.org/10.1371/journal.pmed.1000097>
- Newman, M. et Gough, D. (2020). Systematic reviews in educational research: Methodology, perspectives and application. Dans O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond et K. Buntins (dir.), *Systematic Reviews in Educational Research* (p. 3-22). Springer Fachmedien Wiesbaden. doi: https://doi.org/10.1007/978-3-658-27602-7_1

- Norris, N. (1997). Error, bias and validity in qualitative research. *Educational Action Research*, 5(1), 172-176. doi: <https://doi.org/10.1080/09650799700200020>
- Oakley, A. (2002). Social science and evidence-based everything: The case of education. *Educational Review*, 54(3), 277-286. doi: <https://doi.org/10.1080/0013191022000016329>
- Ouzzani, M., Hammady, H., Fedorowicz, Z. et Elmagarmid, A. (2016). Rayyan. A web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. doi: <https://doi.org/10.1186/s13643-016-0384-4>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. doi: <https://doi.org/10.1136/bmj.n160>
- Pellegrini, M., Lake, C., Neitzel, A. et Slavin, R. E. (2021). Effective programs in elementary mathematics: A meta-analysis. *AERA Open*, 7, 1-29. doi: <https://doi.org/10.1177/2332858420986211>
- Pollock, A. et Berge, E. (2018). How to do a systematic review. *International Journal of Stroke*, 13(2), 138-156. doi: <https://doi.org/10.1177/1747493017743796>
- Regnaud, J.-P. et Remondière, R. (2018). Protocole d'une revue systématique de la littérature des recommandations de bonnes pratiques de kinésithérapie applicables en France. *Kinésithérapie, la Revue*, 18(202), 9-15. doi: <https://doi.org/10.1016/j.kine.2018.05.013>
- Rey, O. (2006). Qu'est-ce qu'une «bonne» recherche en éducation? *Institut français de l'éducation*.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (éd. revue). Newbury Park, CA: Sage Publications.
- Saussez, F. et Lessard, C. (2009). Entre orthodoxie et pluralisme, les enjeux de l'éducation basée sur la preuve. *Revue française de pédagogie*, 168, 111-136. doi: <https://doi.org/10.4000/rfp.1804>
- Schöpfel, J. (2012). Vers une nouvelle définition de la littérature grise. *Cahiers de la documentation/Bladen voor Dokumentatie*, 3, 14-24.
- Slavin, R. E. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*, 48(1), 9-18. doi: [https://doi.org/10.1016/0895-4356\(94\)00097-A](https://doi.org/10.1016/0895-4356(94)00097-A)
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15-21. doi: <https://doi.org/10.3102/0013189X031007015>
- Slavin, R. E. (2008). Perspectives on evidence-based research in education. What works? Issues in synthesizing educational program evaluations. *Educational Researcher*, 37(1), 5-14. doi: <https://doi.org/10.3102/0013189X08314117>
- Slavin, R. E., Lake, C. et Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79(2), 839-911. doi: <https://doi.org/10.3102/0034654308330968>

- Swanson, E., McCulley, L. V., Osman, D. J., Lewis, N. S. et Solis, M. (2017). The effect of team-based learning on content knowledge: A meta-analysis. *Active Learning in Higher Education*. doi: <https://doi.org/10.1177/1469787417731201>
- Tant, M. et Watelain, É. (2014). Attitudes des enseignants d'éducation physique envers l'inclusion d'un élève en situation de handicap. Une revue systématique de la littérature (1975-2015). *Staps*, 106(4), 37-53. doi: <https://doi.org/10.3917/sta.106.0037>
- The Cochrane Collaboration (2020). *Data collection form for intervention reviews: RCTs and non-RCTs*. Repéré à <https://dplp.cochrane.org/data-extraction-forms>
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237-246. doi: <https://doi.org/10.1177/1098214005283748>
- Thomas, J. et Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45. doi: <https://doi.org/10.1186/1471-2288-8-45>
- Tondeur, J., van Braak, J., Ertmer, P. A. et Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development*, 65(3), 555-575. doi: <https://doi.org/10.1007/s11423-016-9481-2>
- Tripney, J., Roulstone, A., Vigurs, C., Hogrebe, N., Schmidt, E. et Stewart, R. (2015). Interventions to improve the labour market situation of adults with physical and/or sensory disabilities in low- and middle-income countries: A systematic review. *Campbell Systematic Reviews*, 11(1), 1-127. doi: <https://doi.org/10.4073/csr.2015.20>
- Vlachopoulos, D. et Makri, A. (2017). The effect of games and simulations on higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education*, 14(1), 22.
- Zaugg, V., Savoldelli, V., Sabatier, B. et Durieux, P. (2014). Améliorer les pratiques et l'organisation des soins: méthodologie des revues systématiques. *Santé publique*, 26(5), 655-667.