

**Revue systématique sur les pratiques évaluatives probantes en
évaluation formative à l'enseignement primaire et secondaire**
**Systematic review of evidence-based formative assessment
practices in elementary and secondary teaching**
**Revisión sistemática de las prácticas evaluativas probatorias
en evaluación formativa para la enseñanza primaria y
secundaria**

Noaga Dieudonné Kaboré, Éric Frenette and Marie-Hélène Hébert

Volume 48, Number 1, 2022

URI: <https://id.erudit.org/iderudit/1096356ar>

DOI: <https://doi.org/10.7202/1096356ar>

[See table of contents](#)

Publisher(s)

Revue des sciences de l'éducation

ISSN

0318-479X (print)

1705-0065 (digital)

[Explore this journal](#)

Cite this article

Kaboré, N. D., Frenette, É. & Hébert, M.-H. (2022). Revue systématique sur les pratiques évaluatives probantes en évaluation formative à l'enseignement primaire et secondaire. *Revue des sciences de l'éducation*, 48(1).
<https://doi.org/10.7202/1096356ar>

Article abstract

The multiple conceptualizations of learning assessment have led to a wide variety of practices, and it is difficult to determine which ones promote student learning. The purpose of this research is to attempt to identify best practices in formative assessment in the context of elementary and secondary education. A systematic review, which identified and analyzed 35 articles, revealed five practices that are recognized as being effective in promoting student learning in the classroom. Two other practices appear promising but require further research. Implications for policies, practices and research are discussed.

Articles

Revue systématique sur les pratiques évaluatives probantes en évaluation formative à l'enseignement primaire et secondaire

Systematic review of evidence-based formative assessment practices in elementary and secondary teaching

Revisión sistemática de las prácticas evaluativas probatorias en evaluación formativa para la enseñanza primaria y secundaria

Contribution 60 %

Noaga Dieudonné KABORÉ

Étudiant à la maîtrise, Université Laval

noaga-dieudonne.kabore.1@ulaval.ca

Contribution 25 %

Éric FRENETTE

Professeur, Université Laval

eric.frenette@fse.ulaval.ca

Contribution 15 %

Marie-Hélène HÉBERT

Professeure, Université TÉLUQ

marie-helene.hebert@teluq.ca

Note éditoriale

Ce texte a été révisé par : Charles-Étienne Tremblay

Résumé

Les multiples conceptualisations de l'évaluation des apprentissages ont conduit à une grande variété de pratiques et il est difficile de déterminer celles permettant de favoriser les apprentissages des élèves. Le but de cette recherche consiste à tenter d'identifier les meilleures pratiques en évaluation formative dans le contexte de l'enseignement primaire et secondaire. Une revue systématique ayant permis d'identifier et d'analyser 35 articles révèle cinq pratiques reconnues probantes pour favoriser les apprentissages des élèves en situation de classe. Deux autres pratiques semblent prometteuses, mais nécessiteront davantage de recherches. Les implications pour les politiques, les pratiques et la recherche sont discutées.

Mots-clés: pratiques évaluatives; évaluation des apprentissages; évaluation formative; données probantes; enseignement primaire et secondaire

Abstract

The multiple conceptualizations of learning assessment have led to a wide variety of practices, and it is difficult to determine which ones promote student learning. The purpose of this research is to attempt to identify best practices in formative assessment in the context of elementary and secondary education. A systematic review, which identified and analyzed 35 articles, revealed five practices that are recognized as being effective in promoting student learning in the classroom. Two other practices appear promising but require further research. Implications for policies, practices and research are discussed.

Keywords: assessment practices; assessment of learning; formative assessment; evidence; elementary and secondary education

Resumen

Las múltiples conceptualizaciones de la evaluación de los aprendizajes han llevado a una gran variedad de prácticas y es difícil determinar cuáles permiten favorecer los aprendizajes de los alumnos. El objetivo de esta investigación consiste en intentar identificar las mejores prácticas en evaluación formativa en el contexto de la enseñanza primaria y secundaria. Nuestro análisis sistemático ha permitido identificar y analizar 35 artículos, revelando cinco prácticas reconocidas como probatorias para favorecer los aprendizajes de los alumnos en clase. Otras dos prácticas parecen prometedoras, pero necesitarán más investigaciones. Finalmente, discutimos las implicaciones para las políticas, las prácticas y la investigación.

Palabras clave: prácticas evaluativas; evaluación de los aprendizajes; evaluación formativa; datos probatorios; enseñanza primaria y secundaria

1. Introduction

Bien que les systèmes éducatifs et les enseignants¹ disposent d'une variété d'outils pour évaluer les apprentissages, la question de l'efficacité des pratiques évaluatives soulève toujours des interrogations, et ce, à tous les niveaux d'enseignement dans les milieux scolaires. Sur le plan de la recherche en éducation, il y a une tendance croissante des études visant à comprendre les activités qui soutiennent l'apprentissage (Wiliam, 2011). Cette tendance est soutenue depuis une dizaine d'années par le recours aux données probantes. Cette perspective de recherche a émergé dans les débats entre les sciences, les pratiques et les politiques après le constat qu'il existe des lacunes entre les savoirs explorés en formation et les savoirs créés par la recherche et auxquels les enseignants font appel dans leurs pratiques de classe (Rey, 2014 ; Slavin, 2020).

Quant aux pratiques évaluatives qui contribuent à bonifier les apprentissages, une hypothèse adoptée presque à l'unanimité dans de nombreuses recherches veut que les activités associées à l'évaluation formative dans les classes fournissent des informations sur le processus d'apprentissage que les enseignants peuvent utiliser pour prendre des décisions pédagogiques afin de bonifier les compétences des élèves (Black et Wiliam, 1998). Cependant, l'évaluation formative au coeur de la régulation des processus d'apprentissage est ouverte à une variété d'interprétations et de pratiques qui ne présentent pas nécessairement toutes les caractéristiques identifiées comme facilitant les progrès des élèves vers un apprentissage ultérieur (Wiliam, 2011). Dans ce contexte, s'assurer quelles sont les pratiques de l'évaluation formative qui aident les élèves à mieux apprendre, demeure une énigme selon diverses études (Dunn et Mulvenon, 2009 ; Rust, O'Donovan et Price, 2005 ; Stiggins, 2002).

Si, de façon générale, les pratiques évaluatives, et plus encore celles formatives, constituent des dispositifs peu étudiés (Clanet et Talbot, 2012 ; Larose, Couturier, Bédard et Charette, 2011), la perspective de recourir aux données probantes dans les recherches est une préoccupation récente en éducation (Slavin, 2020). En évaluation des apprentissages, le recours aux données probantes répond à la nécessité de fournir aux praticiens et aux décideurs un nouvel éclairage sur la question cruciale des meilleures pratiques évaluatives.

Le but de la présente étude consiste à documenter les recherches scientifiques ayant identifié les pratiques en évaluation formative qui favorisent l'apprentissage des élèves dans les contextes de l'enseignement primaire et secondaire. Plus spécifiquement, il s'agit de répertorier, à partir d'une revue systématique, les meilleures pratiques de l'évaluation formative prouvées efficaces dans les écrits scientifiques pour aider les élèves à bonifier leurs apprentissages. Les pratiques évaluatives, l'évaluation des apprentissages et les données probantes sont d'abord définies dans le cadre conceptuel. Puis suivent les aspects méthodologiques de la recherche, la synthèse des articles retenus pour l'analyse, la présentation des résultats et la discussion. Les implications et limites de la recherche sont mentionnées avant la conclusion.

2. Cadre conceptuel

2.1 Les pratiques évaluatives

Les pratiques évaluatives font partie du large concept des pratiques enseignantes. Elles renvoient au « savoir évaluer » des enseignants, c'est-à-dire au système de connaissances rationnelles à la base duquel ces derniers tiennent la mise en oeuvre de l'activité d'évaluation des apprentissages en classe (Deaudelin, Desjardins, Dezutter, Thomas, Corriveau, Lavoie, Bousadra et Hébert, 2007 ;

Morrisette et Nadeau, 2011). Les pratiques évaluatives regroupent divers aspects de la situation d'évaluation, allant de la conception à la sélection des instruments d'évaluation et à leur mise en oeuvre. En tant que telles, les pratiques évaluatives sont des manières concrètes plus ou moins formalisées de penser, d'agir, d'utiliser les outils, les procédures, les langages, les codes et normes et les traces qui se rapportent à l'évaluation des apprentissages (Mottier Lopez, 2015).

En contexte de salles de classe, les pratiques évaluatives sont influencées par la vision de l'enseignement et de l'apprentissage d'une part et par la vision que l'enseignant a de son propre rôle d'autre part (Black et Wiliam, 1998). En cela, les pratiques évaluatives sont étroitement liées aux politiques éducatives qui en définissent et instaurent les valeurs fondamentales ainsi que les décisions concernant les normes et procédures en la matière (Broadfoot et Black, 2004).

2.2 L'évaluation des apprentissages

Le concept d'évaluation des apprentissages est défini diversement dans les écrits. À l'échelle de la classe, l'évaluation est une activité qui vise à recueillir et à interpréter les informations sur les apprentissages scolaires d'un élève ou d'un groupe d'élèves afin de prendre une décision. Vue du côté de l'enseignant, elle permet de juger de l'efficacité de l'acte pédagogique. Au niveau de l'élève, l'évaluation permet de déterminer l'atteinte ou non des objectifs d'apprentissages poursuivis (Laurier, Tousignant et Morissette, 2005). Plus largement, l'évaluation des apprentissages inclut l'ensemble des activités et méthodes utilisées pour obtenir les données sur les connaissances, compétences et comportements des élèves (ministère de l'Éducation du Québec, 2003). Selon les objectifs, les enseignants recourent à des instruments d'évaluation plus ou moins formels ou fréquents et impliquant plus ou moins les élèves (Mottier Lopez, 2015).

Dans les recherches, deux principales caractéristiques ressortent des conceptualisations de l'évaluation des apprentissages scolaires. Ces caractéristiques distinguent l'évaluation de l'apprentissage (*assessment of learning*) de l'évaluation pour l'apprentissage (*assessment for learning*). L'évaluation de l'apprentissage se rapporte à la finalité pour laquelle l'évaluation est effectuée. Elle fait référence à l'évaluation sommative ou certificative, précisément conçue pour servir les objectifs de responsabilité, de classement ou de certification du niveau de compétence des élèves. L'évaluation pour l'apprentissage fait référence à l'évaluation formative. L'une des caractéristiques essentielles de l'évaluation formative réside dans le fait que les preuves des acquisitions des élèves sont interprétées pour déclencher un ensemble d'actions et de pratiques au niveau des enseignants et des élèves qui visent à produire des apprentissages meilleurs (Black et Wiliam, 1998 ; Wiliam, 2011).

Il n'existe pas de consensus sur les composantes de l'évaluation formative (Dunn et Mulvenon, 2009). Cependant, six composantes sont répertoriées dans le rapport de l'Organisation de coopération et de développement économiques (2005). Il s'agit de 1) l'instauration d'une culture de classe encourageant l'interaction et l'utilisation d'outils d'évaluation ; 2) la définition d'objectifs d'apprentissage et du suivi des progrès individuels des élèves ; 3) l'utilisation de méthodes d'enseignement variées ; 4) le recours à des méthodes d'évaluation diversifiées ; 5) garantir une rétroaction sur les performances de l'élève et l'adaptation de l'enseignement en fonction des besoins des élèves et 6) l'implication active des élèves dans le processus d'évaluation. En tant que processus de recherche et d'interprétation d'informations à utiliser par les enseignants pour déterminer où en sont les élèves dans leurs apprentissages afin de prendre des décisions sur ce qui doit être fait pour les aider à y parvenir au mieux, les pratiques de l'évaluation formative nécessitent des recherches probantes (Andersson et Palm, 2017 ; Wiliam, 2011 ; Yorke, 2003).

2.3 Les données probantes en évaluation des apprentissages

Le concept de données probantes (*evidence-based* en anglais) est utilisé dans le sens large des pratiques fondées sur les preuves. La traduction du concept en français est approximative et différentes terminologies sont utilisées selon les domaines et les publications : politiques basées sur les résultats de recherche, pratiques efficaces, ou encore, pratiques ou recherches basées sur les preuves (Rey, 2014 ; Slavin, 2020). Les données probantes font référence à la causalité et sont utilisées pour montrer qu'une chose en rend une autre plus évidente (McQueen et Anderson, 2000). Les pratiques fondées sur les données probantes utilisent des connaissances probabilistes validées par des preuves, contrairement aux approches qui se basent sur la tradition, les conventions, les croyances ou les données non scientifiques (Bissonnette, Gauthier et Péladeau, 2010).

Les recherches montrent que le débat sur le sujet des données probantes était au départ un débat académique sur des questions épistémologiques. Le concept a été utilisé d'abord en médecine avant d'être introduit graduellement dans diverses disciplines (McQueen et Anderson, 2000). Par la suite, les données probantes ont été liées à des questions pragmatiques, notamment pour résoudre des problèmes et prendre des décisions en connaissance de cause. L'identification des données probantes sur une pratique permet d'appliquer une démarche scientifique pour identifier les universaux qui caractérisent l'exercice d'une profession, c'est-à-dire le « que faire » individuel ou collectif (Larose et coll., 2011).

Au point de vue de la méthodologie, les recherches basées sur les données probantes reposent sur des procédés scientifiques spécifiques. Elles recourent aux revues systématiques, aux méta-analyses et aux démarches expérimentales ou quasi expérimentales. Slavin (2020) distingue trois niveaux de preuves : les preuves solides (provenant d'au moins une étude expérimentale), les preuves modérées (provenant d'au moins une étude quasi expérimentale) et les preuves prometteuses provenant d'au moins une étude corrélacionnelle.

Les données probantes en enseignement répondent donc au besoin de constituer un corpus de données scientifiques homologuées par la communauté scientifique attestant ou remettant en cause la validité, le caractère opératoire, relatif ou absolu des connaissances disciplinaires et des savoirs de sens commun propres à la profession enseignante (Larose et coll., 2011). Identifier les pratiques évaluatives probantes, c'est faire connaître ce qui fonctionne efficacement en matière d'évaluation, c'est-à-dire des pratiques évaluatives qui, d'un point de vue éducatif et pédagogique, peuvent améliorer l'expérience d'apprentissage et influencer positivement les rendements des élèves (Organisation de coopération et de développement économiques, 2005). La présente recherche s'intéresse aux données probantes en évaluation formative.

3. Méthodologie

Dans cette section, la méthodologie de la recherche est présentée. L'approche des revues systématiques est retenue pour cette étude. Elle permet d'utiliser une méthode d'identification, d'évaluation et de synthèse des principales connaissances existantes dans un domaine visé suivant une démarche rigoureuse et reproductible (Landry, Becheikh, Amara, Ziam, Idrissi et Castonguay, 2008). Largement utilisées par plusieurs auteurs, les approches des revues systématiques sont plus transparentes, structurées et complètes et leurs résultats sont jugés plus fiables en termes de pertinence, d'utilité ou d'efficacité qu'une revue de littérature (Bearman, Smith, Carbone, Slade, Baik, Hughes-Warrington et Neumann, 2012 ; Dagenais, Martin et Renaud, 2013 ; Munn, Stern, Aromataris, Lockwood et Jordan, 2018).

La présente étude recourt spécifiquement à l'approche méthodologique de Gough (2007) qui utilise une démarche en neuf étapes : 1) la formulation des questions/objectifs de recherche ; 2) la définition des critères d'inclusion et d'exclusion ; 3) l'articulation de la stratégie de recherche documentaire ; 4) la vérification des critères d'inclusion et d'exclusion des articles ; 5) le report des résultats de la stratégie de recherche, 6) l'extraction des données pertinentes pour la recherche ; 7) l'évaluation de la qualité méthodologique des articles retenus ; 8) la synthèse quantitative ou qualitative des résultats communs des articles ; 9) la présentation synthétique des résultats et les conclusions de la recherche.

À l'étape 1, le but de l'étude a préalablement été défini dans l'introduction. À l'étape 2, six critères d'éligibilité des études ont été fixés. Ces critères sont relatifs a) au titre, b) à la période de publication, c) à la langue de publication, d) au résumé, e) au texte de l'article et f) au contexte scolaire. Sur le plan du titre, l'article devrait contenir au moins un des trois concepts de la recherche ou leurs synonymes (tableau 1). Quant aux critères d'inclusion/exclusion, les articles devraient être publiés dans une revue scientifique avec comité de lecture au cours des 22 dernières années, soit entre 1998 et 2020. Ces articles devraient être en français ou en anglais et leur texte intégral devait être accessible. Les articles théoriques et professionnels, les rapports gouvernementaux, les communications de conférences et les sections de livres ont été exclus. Les recherches documentaires acceptées devraient être des revues systématiques, des méta-analyses ou des critiques de méta-analyses.

Dans les caractéristiques du résumé, les articles devraient être de devis quantitatif, qualitatif ou mixte, tout en faisant référence à un des trois concepts à l'étude ou leurs synonymes. Les objectifs, le contexte de l'étude, les méthodes d'analyse ainsi que les résultats devraient être clairement précisés. Le dernier critère fait référence au contexte de l'évaluation des apprentissages dans l'enseignement primaire et postprimaire (ou secondaire) (tableau 1). Les articles ayant utilisé le terme « collègue » dans le sens de l'enseignement préuniversitaire ou universitaire n'ont pas été retenus.

À l'étape 3 de la démarche, une bibliothécaire professionnelle de l'université d'attache a été consultée pour aider à l'élaboration de la stratégie de recherche documentaire. La chaîne de recherche a été formulée et une liste des concepts opératoires (mots clés) et leurs synonymes a été dressée en vue des recherches sur les bases de données. Les synonymes utilisés proviennent de la littérature et du thésaurus (vocabulaire contrôlé du dictionnaire) (tableau 1). Les concepts en anglais et leurs synonymes ont été utilisés pour la recherche dans trois bases de données (Education Source, ERIC, PsycINFO).

Tableau 1

Les concepts opératoires et leurs synonymes pour la recherche documentaire

Concepts en français et leurs synonymes	Concepts en anglais et leurs synonymes
Concept 1 – Pratiques évaluatives	
« Évaluation des apprentissages ; pratiques évaluatives ; méthodes d'évaluation ; évaluation formative ; portfolio en éducation ; évaluation en éducation ; évaluation des curricula »	« Evaluation ; assessment ; evaluation practice ; evaluation methods ; formative evaluation ; formative assessment ; assessment for learning ; portfolio in education ; educational evaluation ; education-assessment ; curriculum evaluation »
Concept 2 – Enseignement primaire et secondaire	
« Enseignement primaire ; enseignement élémentaire ; enseignement secondaire ; lycée, collège »	« Primary education ; elementary school ; primary education ; secondary school ; primary school teaching ; elementary school teaching ; elementary education ; high school teaching ; secondary education ; junior high school teaching »
Concept 3 – Données probantes	
« Pratiques probantes en éducation ; données probantes en éducation ; meilleures pratiques évaluatives en éducation »	« Evidence-based education ; evidence-based practices ; best practices »

La recherche documentaire a été réalisée entre octobre et décembre 2020 (figure 1). Sept bases de données contenant la plupart des publications en éducation ont été consultées. Ces bases sont Education Source, ERIC, PsycINFO pour les bases de données en anglais et Ariane, Erudit, Cairn.info, les bases bibliographiques Pascal et Francis en français. Des recherches complémentaires sur Google Scholar et un examen de références des articles répertoriés ont été effectués. La synthèse de la chaîne de recherche des articles est présentée dans la figure 1.

À l'étape 4, 2741 articles trouvés suivant la stratégie de recherche ont été exportés vers le logiciel Endnote X9 et soumis à la vérification des critères d'inclusion et d'exclusion. La vérification a permis d'éliminer 1614 doublons. Le reste des 1127 articles a été soumis à l'application du critère relatif à la langue permettant de supprimer 104 articles. Au total, 1023 articles ont été soumis à l'application du critère relatif à l'année de publication et à la publication dans une revue scientifique avec comité de lecture. L'application de ces critères a permis de supprimer 961 articles et d'en sélectionner 62. L'examen complet de ces articles a permis d'évaluer la pertinence de leur qualité relativement aux critères d'inclusion liés au titre, au résumé, au contexte scolaire et à la disponibilité des textes. Il est apparu que 30 articles ne correspondaient pas aux critères et ont été supprimés lors du report des résultats et de l'extraction des données aux étapes 5 et 6. Des articles supprimés, 14 avaient des questions de recherche non pertinentes, sept n'avaient pas précisé la méthodologie utilisée et neuf autres études ont porté sur un échantillon universitaire.

Enfin, l'examen des références des 32 articles retenus a permis d'identifier trois articles supplémentaires qui satisfaisaient tous les critères d'inclusion. En définitive, la recherche documentaire a permis d'identifier 35 articles pertinents portant sur les pratiques d'évaluation formative des apprentissages. Ces articles ont été analysés en suivant les critères d'une évaluation fondée sur des données probantes. Le contrôle de la qualité méthodologique des articles retenus (clarté des objectifs/questions de recherche, méthodes d'analyse appropriées, qualité de la présentation des résultats) effectué à l'étape 7 montre que ces études se conforment aux exigences

méthodologiques fixées. Les articles retenus ont été lus intégralement et les résultats pertinents ont été enregistrés manuellement pour la synthèse à l'étape 8 et la présentation synthétique des résultats et les conclusions de la recherche à l'étape 9.

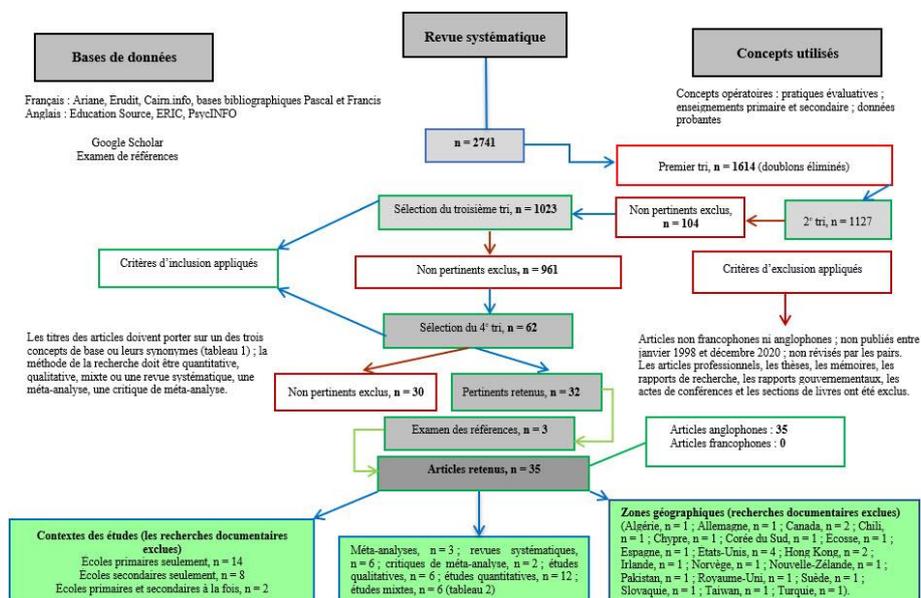
4. Caractéristiques des articles retenus

Six types de recherches répartis en deux catégories composent les articles retenus pour l'analyse. Des recherches documentaires composées de revues systématiques (n = 6), de méta-analyses (n = 3) et de critiques de méta-analyses (n = 2) ayant examiné des études concernant les pratiques de l'évaluation formative forment la première catégorie. La deuxième catégorie est composée d'études réalisées en contexte de classe. Des recherches quantitatives (n = 12), qualitatives (n = 6) et mixtes (n = 6) sur les pratiques de l'évaluation formative au primaire ou au secondaire ont été retenues dans cette catégorie. Les détails sont présentés dans l'annexe.

Toutes les études sont de langue anglaise. Selon le contexte géographique (excluant revues systématiques, méta-analyses et critiques de méta-analyse), 24 recherches ont été réalisées dans 19 pays. Neuf études ont été menées en Europe (Allemagne, Chypre, Écosse, Espagne, Irlande, Norvège ; Royaume-Uni, Slovaquie, Suède). Sept études ont été réalisées en Amérique (Canada, Chili et États-Unis) et six études relèvent du contexte asiatique (Corée du Sud ; Hong Kong ; Pakistan ; Taiwan ; Turquie). Une étude a été menée en Afrique du Nord (Algérie) autant qu'en Océanie (Nouvelle-Zélande).

Figure 1

Synthèse de la chaîne de recherche des articles



Alors que les recherches documentaires ont analysé des études ayant porté sur les pratiques de l'évaluation formative dans les contextes scolaires du primaire et du secondaire sans distinction, 14 études réalisées en contexte de classe se sont déroulées au primaire, deux études ont eu lieu à la fois au primaire et au secondaire et huit études se sont déroulées exclusivement au secondaire. Les disciplines couvertes par l'ensemble des études sont les mathématiques (n = 5), le français (n = 2), la lecture/écriture (n = 3), l'anglais (n = 4), l'informatique (n = 1), la biologie (n = 2), la chimie (n = 1), l'éducation physique (n = 1) et les études sociales (n = 1). Deux études ont concerné plus d'une discipline à la fois (lecture/français, mathématiques, biologie, anglais, études sociales). Sept études n'ont pas précisé les disciplines.

Sur le plan de la méthodologie, les revues systématiques ont analysé entre 17 et 39 articles. Une étude a recouru à une approche méta-analytique à effets aléatoires tandis que les deux autres méta-analyses ont utilisé l'analyse de contenu. Les méta-analyses ont respectivement porté sur l'examen de 13 à 40 études empiriques. Les critiques de méta-analyse ont examiné deux méta-analyses.

Les recherches quantitatives ont utilisé un devis expérimental (n = 6) et quasi expérimental (n = 6). Les instruments de collecte des données sont essentiellement les questionnaires et les tests. Les échantillons aléatoires des études variaient de n = 55 à n = 13 567 pour les élèves et de n = 1 à n = 906 pour les enseignants repartis dans des écoles (entre n = 1 à n = 360). Les modèles d'analyse des données sont les statistiques descriptives, les corrélations, le test *t*, l'ANOVA, l'ANCOVA, les régressions linéaires et les tests chi carré.

Les études qualitatives ont emprunté le modèle d'étude de cas. Les entrevues, les observations directes, les prises de notes et les artéfacts ont été les méthodes de collecte des données. Les échantillons variaient de 2 à 118 pour les enseignants et de 21 à 305 pour les élèves. L'analyse par thèmes émergents, l'analyse de contenu et la comparaison constante ont été les moyens d'analyse des données qualitatives.

Les approches quantitatives des études mixtes ont utilisé un devis quasi expérimental. Les volets qualitatifs ont adopté le modèle d'étude de cas. Les moyens de collecte des données étaient des questionnaires et des tests pour les approches quantitatives. Les études de cas ont recouru à des entrevues semi-dirigées, à l'observation directe et l'analyse de données documentaires existantes. Les échantillons dans ces études varient entre 4 et 162 participants pour les enseignants et entre 100 et 375 pour les élèves. Alors que les analyses quantitatives des recherches mixtes ont recouru aux statistiques descriptives combinées au test *t* et au test ANOVA, le volet qualitatif a utilisé l'analyse par thèmes émergents et l'analyse de contenu.

5. Résultats

Le but de cette recherche consiste à tenter d'identifier les meilleures pratiques en évaluation formative prouvées dans le contexte de l'enseignement primaire et secondaire. Sept pratiques de l'évaluation formative reconnues comme meilleures ressortent des résultats des articles soumis à l'analyse. Le regroupement des différentes études par thématique en lien avec ces pratiques est fait au tableau 2 ci-dessous.

5.1 Les meilleures pratiques prouvées en évaluation formative

Cette section présente les résultats essentiels concernant les pratiques décrites dans les études comme les meilleures pratiques prouvées en évaluation formative. Les résultats d'analyse détaillés incluant les articles et leur méthodologie sont en annexe.

5.1.1 L'autoévaluation et l'évaluation par les pairs

L'autoévaluation est une situation d'évaluation où l'élève est amené à porter un jugement sur ses propres compétences à travers l'établissement d'objectifs et de critères d'appréciation précis. Le processus d'évaluation par les pairs amène l'apprenant à porter un jugement sur le travail d'autrui (Leclerc, 2013). L'autoévaluation et l'évaluation par les pairs sont examinées à travers dix études, dont une revue systématique (Sebba, Crick, Yu, Lawson, Harlen et Durant, 2008), une méta-analyse (Sanchez, Atkinson, Koenka, Moshontz et Cooper, 2017), trois recherches quantitatives (Chang, Liang et Chen, 2013 ; Panadero, Tapia et Huertas, 2012 ; Sadler et Good, 2006), deux recherches qualitatives (Ní Chróinín et Cosgrave, 2013 ; Noonan et Duncan, 2005) et trois recherches mixtes (Burner, 2016 ; Butler et Lee, 2010 ; Vandergrift, 2000). La moitié des études ayant abordé les deux pratiques à la fois, nous décidons de regrouper leurs résultats sous la même rubrique.

Tableau 2

Synthèse des études par thématique en lien avec l'évaluation formative

	Études quantitatives	Études qualitatives	Études mixtes	Méta-analyses	Critiques/méta-analyses	Revue systématiques	Total	Articles (premier auteur et année de publication)
Autoévaluation (seule)	2	1	2	-	-	-	5	(Burner, 2016 ; Butler et coll., 2010 ; Chang et coll., 2013 ; Ní Chróinín et coll., 2013 ; Panadero et coll., 2012)
L'évaluation par les pairs (seule)	-	-	-	-	-	-	-	-
L'autoévaluation et l'évaluation par les pairs (à la fois)	1	1	1	1	-	1	5	(Noonan et coll., 2005 ; Sadler et coll., 2006 ; Sanchez et coll., 2017 ; Sebba et coll., 2008 ; Vandergrift, 2000)
L'autorégulation	1	-	-	-	-	-	1	(Perels et coll., 2009)
L'évaluation par le portfolio	-	2	-	-	-	1	3	(Burner, 2014 ; Mak et coll., 2018 ; McLaren, 2012)
L'utilisation des grilles d'évaluation	-	-	-	-	-	1	1	(Panadero et coll., 2013)
La rétroaction	-	1	1	-	-	-	2	(Lee, 2007 ; Parr et coll., 2010)
La carte conceptuelle	-	-	-	-	-	1	1	(Stevenson et coll., 2017)
La mise en œuvre et les impacts des pratiques de l'évaluation formative sur les élèves et leur apprentissage	8	1	2	2	2	2	17	(Andersson et coll., 2017 ; Antoniou et coll., 2014 ; Babinčáková et coll., 2020 ; Black et coll., 1998 ; Boumediene et coll., 2017 ; Burns et coll., 2010 ; Büyükkarci, 2014 ; Dunn et coll., 2009 ; Heitink et coll., 2016 ; Kingston et coll., 2011 ; Kirton et coll., 2007 ; Leong et coll., 2018 ; McMillan et coll., 2013 ; Phelan et coll., 2011 ; Polly et coll., 2017 ; Saeed et coll., 2018 ; Vásquez et coll., 2017)
Total	12	6	6	3	2	6	35	

La revue systématique (Sebba et coll., 2008) et la méta-analyse (Sanchez et coll., 2017) ont porté sur les effets de l'autoévaluation et de l'évaluation par les pairs sur les apprentissages des élèves. Ayant analysé 26 études empiriques, les résultats de la revue systématique de Sebba et coll. (2008) révèlent que les pratiques de l'autoévaluation et de l'évaluation par les pairs sont des moyens d'autorégulation ayant une incidence positive sur les rendements scolaires des élèves. Leur mise en oeuvre transforme la relation de dépendance des élèves avec l'enseignant en une relation d'interdépendance, tout en augmentant leur estime de soi et leur engagement envers l'apprentissage. Cependant, l'apport de ces deux pratiques diffère selon le sexe, l'origine ethnique ou les résultats scolaires antérieurs des élèves.

Fondés sur l'analyse de 33 articles, les résultats de la méta-analyse effectuée par Sanchez et coll. (2017) démontrent qu'au-delà du développement des compétences autorégulatrices chez les élèves du primaire et du secondaire, l'autoévaluation et l'évaluation par les pairs renforcent leurs capacités de rétention en rendant plus claire leur compréhension de la matière évaluée. Alors que les gains des élèves qui se sont évalués ont été estimés comme étant plus importants dans les études analysées, l'étude conclut que l'autoévaluation et l'évaluation par les pairs sont des pratiques d'évaluation formative efficaces qui aident les élèves à mieux apprendre et à mieux réussir aux tests subséquents.

Les trois études quantitatives (Chang et coll., 2013 ; Panadero et coll., 2012 ; Sadler et Good, 2006) ont porté sur l'autoévaluation et l'évaluation par les pairs dans le contexte de l'enseignement secondaire à Taiwan, à Madrid et aux États-Unis. En plus de vérifier leurs impacts sur l'apprentissage des élèves, deux de ces études (Chang et coll., 2013 ; Sadler et Good, 2006) ont examiné leur validité et leur fiabilité. Les résultats des trois études indiquent que non seulement l'autoévaluation et l'évaluation par les pairs par les élèves sont des méthodes dont la fiabilité et la validité sont appropriées, mais elles développent également les capacités autorégulatrices des élèves tout en les conduisant à de meilleurs apprentissages. Les élèves ayant participé à l'autoévaluation obtiennent de meilleurs résultats aux tests subséquents comparativement à ceux ayant été évalués par leurs pairs dans l'étude de Sadler et Good (2006).

Les potentialités de l'autoévaluation et de l'évaluation par les pairs sont enfin rapportées par les études qualitatives (Ní Chróinín et Cosgrave, 2013 ; Noonan et Duncan, 2005) et mixtes (Burner, 2016 ; Butler et Lee, 2010 ; Vandergrift, 2000) ayant interrogé l'expérience des enseignants et des élèves de leur utilisation de ces pratiques. Les résultats révèlent que, selon les enseignants, l'autoévaluation et l'évaluation par les pairs sont des approches flexibles et adaptées pour la responsabilisation, la motivation et l'autonomie des élèves. Combinées au partage des objectifs d'apprentissage et des critères d'évaluation, ces pratiques aident les élèves à bonifier leurs apprentissages par le biais de la rétroaction. Cependant, les résultats de trois études (Burner, 2016 ; Butler et Lee, 2010 ; Noonan et Duncan, 2005) révèlent que les enseignants ne peuvent pas les utiliser fréquemment dans chaque classe. Ces derniers préfèrent l'évaluation par les pairs en raison de l'immaturation ou de l'incapacité de leurs élèves à s'engager avec précision et équité dans l'autoévaluation.

5.1.2 L'autorégulation

L'autorégulation en évaluation fait référence à l'ensemble des activités d'évaluation qui vise à développer la capacité de l'élève à planifier, surveiller et évaluer son apprentissage à travers ses propres stratégies (Leclerc, 2013 ; Organisation de coopération et de développement économiques, 2005). La pratique de l'autorégulation a été examinée dans une étude quantitative (Perels, Dignath et Schmitz, 2009). Elle a consisté à examiner les effets de la formation des compétences d'autorégulation sur l'apprentissage ainsi que sur la réussite en mathématiques dans le contexte de l'école primaire en Allemagne. Les résultats rapportés des tests ANOVA à mesures répétées révèlent que les élèves (n = 55) des classes expérimentales ont amélioré leurs rendements en mathématiques à la suite d'évaluations formatives soutenues par l'autorégulation. L'étude conclut que le développement de l'autorégulation offre des possibilités d'apprentissage étendues tout en ayant un potentiel d'améliorer les rendements des élèves.

5.1.3 L'évaluation par le portfolio

L'efficacité de l'évaluation par le portfolio est soutenue par les résultats de trois études, dont une revue systématique (Burner, 2014) et deux études qualitatives (Mak et Wong, 2018 ; McLaren, 2012). La revue systématique a étudié le potentiel formatif de l'évaluation par le portfolio dans les cours de rédaction à l'école primaire. Les résultats basés sur l'analyse des 39 articles révèlent qu'en facilitant la rétroaction et la révision des textes par les élèves, l'évaluation par le portfolio est une pratique d'évaluation exempte d'anxiété qui impacte positivement à la fois les processus d'apprentissage et les rendements des élèves.

Mak et Wong (2018) ont examiné comment les enseignants du primaire encouragent l'autorégulation des élèves du primaire (n = 69) par le biais de l'évaluation par le portfolio dans le contexte de Hong Kong. Quant à la recherche de McLaren (2012), elle a analysé les potentialités d'une approche *e-scape* (*e-solutions for creative assessment in portfolio environment*) pour soutenir l'évaluation de l'apprentissage. Les résultats des deux études démontrent que l'évaluation par le

portfolio est un moyen pour le diagnostic et la rétroaction en évaluation formative. En cela, les études concluent que le portfolio est une activité de responsabilisation des élèves qui contribue à promouvoir l'apprentissage autorégulé à travers la participation des élèves à la réflexion, ainsi que le développement de leur autonomie et de leurs capacités en autoévaluation.

5.1.4 L'utilisation des grilles d'évaluation

L'efficacité de l'utilisation des grilles d'évaluation pour soutenir l'apprentissage a été abordée dans une revue systématique (Panadero et Jonsson, 2013). Portant sur 21 articles ayant examiné si et comment les grilles d'évaluation ont un impact sur l'apprentissage des élèves, les résultats de cette revue systématique révèlent que l'élaboration et le partage des critères d'évaluation ainsi que leur utilisation avec les élèves accroissent la transparence de l'évaluation et aident au processus de rétroaction tout en améliorant l'autonomie des élèves. Cependant, l'efficacité des grilles d'évaluation dépend de la qualité des indicateurs d'apprentissage et de la participation directe des élèves à leur élaboration et à leur interprétation.

5.1.5 La rétroaction

La rétroaction ressort de façon générale dans plusieurs études comme une composante cruciale de la remédiation dont les fonctions aident les élèves à apprendre. Ses formats et impacts ont été spécifiquement examinés dans une étude qualitative (Lee, 2007) et une étude mixte (Parr et Timperley, 2010). La recherche qualitative a analysé dans le contexte de Hong Kong la nature et les fonctions de la rétroaction des enseignants ($n = 26$) afin de déterminer dans quelle mesure elle permet de stimuler l'évaluation formative et les apprentissages des élèves du secondaire ($n = 174$). L'analyse par thèmes émergents des textes des élèves révèle que la rétroaction a eu un effet positif ou négatif selon sa nature, sa qualité et sa variété. Les élèves préfèrent les rétroactions orales interactives par rapport à celles écrites et les rétroactions constructives positives aux rétroactions négatives et directives. L'étude de Parr et Timperley (2010) a visé à définir les caractéristiques de la rétroaction de qualité chez des enseignants du primaire ($n = 59$) et ses impacts sur l'apprentissage des élèves ($n = 375$) dans le contexte de l'école primaire en Nouvelle-Zélande. Les résultats montrent qu'une rétroaction directive faite de commentaires négatifs et non ciblés est frustrante pour les élèves, contrairement aux rétroactions orales et positives et aux discussions interactives qui sont jugées comme deux types de rétroaction plus utiles. Sur le plan des effets sur l'apprentissage, les résultats démontrent qu'il y a une corrélation significative moyennement élevée entre la qualité de la rétroaction fournie par les enseignants et l'étendue moyenne des gains en écriture chez les élèves ($r = 0,69$). L'étude conclut que la rétroaction, dépendamment de ses qualités, favorise l'efficacité de l'évaluation formative et contribue à développer les compétences des élèves en écriture.

5.1.6 La carte conceptuelle

La carte conceptuelle ou cartographie conceptuelle est un outil représentant de façon structurée un ensemble de concepts reliés sémantiquement dont les relations de construction sont décrites par des mots (Farza, 2018). Une revue systématique (Stevenson, Hartmeyer et Bentsen, 2017) a examiné comment les technologies de cartographie conceptuelle sont un outil stimulant l'apprentissage autorégulé chez les élèves du primaire et du secondaire. L'analyse basée sur 17 articles révèle que les technologies de cartographie conceptuelle influencent positivement l'apprentissage autorégulé tout en améliorant les compétences métacognitives et la motivation des élèves à divers degrés.

5.2 La mise en oeuvre de l'évaluation formative

Les impacts de la mise en oeuvre de l'évaluation formative en tant qu'unité de pratiques intégrées sont examinés dans près de la moitié des études (17 sur 35). Au niveau des recherches documentaires, deux revues systématiques (Heitink, Van der Kleij, Veldkamp, Schildkamp et Kippers, 2016 ; Leong, Ismail, Costa et Tan, 2018) et deux méta-analyses (Black et Wiliam, 1998 ; Kingston et Nash, 2011) ont examiné cette question. Les revues systématiques ont porté sur la mise en oeuvre de l'évaluation formative dans les contextes de l'Asie de l'Est (Leong et coll., 2018) et de l'Australie (Heitink et coll., 2016). Les résultats indiquent que, malgré les conceptualisations divergentes de l'évaluation formative présentées dans les recherches, des preuves indiquent, dans les deux revues systématiques, que les pratiques de l'évaluation formative aident les élèves à mieux apprendre.

Cependant, les effets positifs des pratiques de l'évaluation formative sont mis en lien avec l'efficacité de son implémentation. L'efficacité des pratiques en évaluation formative est dépendante de la qualité de la rétroaction fournie aux élèves, de l'implication de ces derniers dans l'élaboration et le partage des objectifs et critères d'évaluation, de l'amélioration des compétences des enseignants et des élèves à l'utilisation des stratégies de l'évaluation formative. À ce niveau, les résultats des deux revues systématiques révèlent unanimement que les impacts positifs des pratiques de l'évaluation formative sur l'apprentissage des élèves sont faibles, limités ou nuls lors de mises en oeuvre inefficaces.

Des facteurs tels le manque de capacités des enseignants et la pression des examens externes sont signalés comme des obstacles majeurs à la mise en oeuvre de l'évaluation formative. Dans certains contextes, des traditions d'apprentissage enracinées faisant de l'enseignant le détenteur ultime des connaissances créent aussi des attitudes, tant chez les élèves que chez les enseignants, qui posent des défis à une démarche collaborative de l'évaluation où les responsabilités de l'apprentissage sont partagées avec les élèves (Leong et coll., 2018). Les deux études suggèrent que l'efficacité de l'évaluation formative exige un équilibre entre les connaissances, compétences, croyances et attitudes aussi bien des enseignants que des élèves.

Les résultats de la méta-analyse de Black et Wiliam (1998) et de celle de Kingston et Nash (2011) ont montré que les pratiques efficaces de l'évaluation formative ont des impacts positifs importants sur l'apprentissage des élèves, quel que soit leur niveau scolaire. Black et Wiliam (1998) incluent, dans les qualités de l'évaluation formative, deux principales caractéristiques, à savoir la diversification des instruments et la rétroaction fournie pour aider les plus faibles. La taille d'effet de l'évaluation formative est comprise entre 0,40 et 0,70 dans l'étude de Black et Wiliam (1998) et dans celle de Kingston et Nash (2011), elle est située entre 0,20 et 0,25. L'évaluation formative serait plus efficace en enseignement de l'anglais qu'en enseignement des mathématiques ou des sciences (taille d'effet moyen respective : 0,32 ; 0,17 et 0,09). L'évaluation formative intervenant après une formation des enseignants et les pratiques soutenues par les systèmes informatisés se révèlent plus efficaces que celles basées sur les approches ordinaires avec une taille de l'effet moyen de 0,30 et 0,28 respectivement en anglais et en mathématiques (Kingston et Nash, 2011).

Les résultats de ces deux méta-analyses sont remis en cause dans les analyses critiques de Dunn et Mulvenon (2009) et McMillan, Venable et Varier (2013). Relevant des problèmes méthodologiques dans ces deux méta-analyses, les examens critiques révèlent que les effets estimés de l'évaluation formative seraient plus faibles que ceux présentés dans les méta-analyses. Selon les auteurs, la certitude que l'utilisation de l'évaluation formative augmente directement les résultats des élèves est une idée répandue par convention. L'inexistence d'une définition et de composantes exactes de l'évaluation formative profite à « des approches méthodologiques suspectes dans les efforts visant à démontrer des effets positifs qui pourraient être attribués aux évaluations formatives » (Dunn et Mulvenon, 2009, p. 1, traduction libre). Les auteurs suggèrent que les généralisations sont

inappropriées et qu'en l'absence de plus de données empiriques, les résultats revendiqués dans les méta-analyses de Black et Wiliam (1998) et de Kingston et Nash (2011) doivent être considérés comme provisoires.

Les résultats de huit études quantitatives (Andersson et Palm, 2017 ; Babinčáková, Ganajová, Sotáková et Bernard, 2020 ; Boumediene et Hamzaoui-Elachachi, 2017 ; Burns, Klingbeil et Ysseldyke, 2010 ; Phelan, Choi, Vendlinski, Baker et Herman, 2011 ; Polly, Wang, Martin, Lambert, Pugalee et Middleton, 2017 ; Saeed, Tahir et Latif, 2018 ; Vásquez, Nussbaum, Sciarresi, Martínez, Barahona et Strasser, 2017) rapportent aussi des preuves de l'efficacité des pratiques de l'évaluation formative dans divers contextes (Algérie, Chili, États-Unis, Pakistan, Suède, Slovaquie). Les résultats des différentes études révèlent des impacts positifs, aussi bien sur les processus d'apprentissage que sur l'amélioration des rendements chez les élèves du groupe expérimental par rapport aux élèves des groupes témoins en français, en lecture/écriture, en orthographe et en mathématique. Cependant, l'étude de Saeed et coll. (2018), dans le contexte pakistanais, indique que la plupart des enseignants (n = 500) des écoles primaires et secondaires reconnaissent les potentialités de l'évaluation formative, mais que ces derniers utilisent plus l'évaluation sommative.

Les résultats des recherches de Burns et coll. (2010), Polly et coll. (2017) et de Vásquez et coll. (2017) qui ont examiné l'utilisation d'un programme d'évaluation formative basée sur la technologie, respectivement dans le contexte de Floride, de la Caroline du Nord et du Chili, rapportent des preuves de l'efficacité de l'évaluation formative assistée par ordinateur. L'étude de Burns et coll. (2010) a porté sur un échantillon de 360 écoles. La recherche de Polly et coll. (2017) a suscité la participation de 906 enseignants et 13 567 élèves. L'étude de Vásquez et coll. (2017) a porté sur un échantillon de 90 élèves. Les résultats issus des différentes analyses montrent une augmentation du niveau de rendements de tous les élèves du groupe expérimental grâce à l'évaluation formative assistée par ordinateur. Il y a un effet positif sur le pourcentage d'élèves qui réussissent aux évaluations sommatives (Burns et coll., 2010). Les études concluent que l'évaluation formative assistée par les technologies est plus efficace pour atteindre de meilleurs apprentissages et rendements.

Enfin, deux études qualitatives (Antoniou et James, 2014 ; Büyükkarci, 2014) et une étude mixte (Kirton, Hallam, Peffer, Robertson et Stobart, 2007) qui ont examiné comment les enseignants et les élèves perçoivent l'efficacité de l'évaluation formative parviennent à des résultats comparables à ceux des études précédentes. Les résultats indiquent que les pratiques formatives sont adoptées sans réserve par les enseignants et les écoles qui jugent que les principales pratiques mises en oeuvre impactent positivement la responsabilisation des élèves, leur motivation, leur autonomie et leurs rendements dans les tests sommatifs. Cependant, bien que les enseignants aient des perceptions positives de l'utilisation des pratiques de l'évaluation formative, leur défaut de compétences et le manque de temps limitent leur mise en oeuvre effective (Antoniou et James, 2014).

6. Discussion

Notre recherche visait à documenter les recherches scientifiques ayant identifié les meilleures pratiques évaluatives en évaluation formative qui favorisent l'apprentissage des élèves dans les contextes de l'enseignement primaire et secondaire à partir d'une revue systématique. La recherche a identifié 35 études probantes révisées par les pairs. Sept pratiques clés de l'évaluation formative sont promues dans les recherches : l'autoévaluation, l'autorégulation, l'évaluation par les pairs, l'évaluation par le portfolio, l'utilisation des grilles d'évaluation, la carte conceptuelle et la rétroaction.

Deux hiérarchies de preuves concernant les pratiques de l'évaluation formative se dégagent des résultats des études : les pratiques confirmées et les pratiques en développement. Les pratiques confirmées sont celles dont les preuves sont supportées par les résultats de plusieurs études d'envergure, dont les revues systématiques, les méta-analyses, les critiques de méta-analyse, les études quantitatives expérimentales ou quasi expérimentales ainsi que les études qualitatives.

Cinq pratiques de l'évaluation formative sont des pratiques probantes confirmées. Il s'agit de l'autoévaluation, de l'évaluation par les pairs, de l'évaluation par le portfolio, de l'utilisation des grilles d'évaluation et de la carte conceptuelle. Ces pratiques confirmées de l'évaluation formative ont une incidence sur les processus d'apprentissages et sur les rendements des élèves. Leur efficacité est discutée à travers plusieurs études probantes, dont les revues systématiques (Burner, 2014 ; Panadero et Jonsson, 2013 ; Sebba et coll., 2008 ; Stevenson et coll., 2017), les méta-analyses (Black et Wiliam, 1998 ; Sanchez et coll., 2017), les études quantitatives (Babinčáková et coll., 2020 ; Chang et coll., 2013 ; Panadero et coll., 2012 ; Sadler et Good, 2006), qualitatives (Lee, 2007 ; Mak et Wong, 2018 ; McLaren, 2012 ; Ní Chróinín et Cosgrave, 2013 ; Noonan et Duncan, 2005) et mixtes (Burner, 2016 ; Butler et Lee, 2010 ; Vandergrift, 2000).

Les pratiques en développement sont celles identifiées par les recherches comme meilleures, mais dont les preuves d'efficacité sont toujours en train d'émerger en regard du nombre limité d'études probantes qui supportent leurs résultats. Elles concernent les pratiques d'autorégulation et les formats des rétroactions. Si l'autorégulation conduit les élèves à de meilleurs apprentissages, le développement des compétences autorégulatrices mis en lien avec la pratique de l'autoévaluation dans les études (Mak et Wong, 2018 ; Panadero et coll., 2012) n'est supporté que par une étude (Perels et coll., 2009).

Quant à la rétroaction, bien que sa pratique ressorte dans plusieurs études comme une condition substantielle de la remédiation et de l'efficacité de l'évaluation formative (Black et Wiliam, 1998 ; Heitink et coll., 2016), la relation entre la qualité ou le format des rétroactions fournies aux élèves et le développement de leurs apprentissages n'est pas discutée dans plusieurs études. Les études qui discutent de l'efficacité de la qualité de la rétroaction à partir de sa nature et de ses variétés (Burner, 2016 ; Lee, 2007 ; Parr et Timperley, 2010) indiquent que les rétroactions normatives et négatives sont inefficaces, contrairement aux rétroactions interactives et positives, reconnues favorables à l'autorégulation des apprentissages. Ces conclusions sont conformes aux résultats des recherches de Grangeat et Lepareur (2019) et d'Hattie et Timperley (2007) qui ont examiné la nature et les conditions de mise en oeuvre de la rétroaction en classe. Ces études notent que la rétroaction n'est pas efficace pour tous les élèves et que l'état actuel des recherches ne présente que des esquisses de modèles de rétroactions à utiliser.

Les impacts positifs de la mise en oeuvre de l'évaluation formative en tant qu'ensemble des pratiques sur les apprentissages des élèves sont largement promus dans la majorité des études retenues. Que ce soit au niveau des revues systématiques (Burner ; 2014 ; Panadero et Jonsson, 2013 ; Sebba et coll., 2008 ; Stevenson et coll., 2017), des méta-analyses (Black et Wiliam, 1998 ; Sanchez et coll., 2017), des études quantitatives (Andersson et Palm, 2017 ; Babinčáková et coll., 2020 ; Boumediene et Hamzaoui-Elachachi, 2017 ; Panadero et coll., 2012 ; Perels et coll., 2009 ; Phelan et coll., 2011 ; Polly et coll., 2017 ; Sadler et Good, 2006), qualitatives (Antonioni et James, 2014) et mixtes (Büyükkarci, 2014 ; Kirton et coll., 2007), des preuves aident à soutenir l'idée que la pratique de l'évaluation formative aide les élèves à mieux apprendre en agissant sur trois domaines en situation d'apprentissage : la motivation à étudier, la prise de conscience par les élèves de leur propre processus d'apprentissage et leur rendement scolaire.

Les revues systématiques et les méta-analyses ont montré que les pratiques d'une évaluation formative de qualité ont des impacts positifs importants sur les apprentissages des élèves, quel que soit leur niveau scolaire, et ce, même si elles doivent être relativisées par les critiques de méta-analyse (McMillan et coll., 2013 ; Dunn et Mulvenon, 2009). Les effets positifs constatés dans les

études quantitatives qui ont examiné l'impact de l'évaluation formative et de ses pratiques sur les apprentissages et les rendements corroborent son rôle crucial dans la promotion de l'apprentissage en classe. Les recherches montrent que, lorsque les enseignants reçoivent des formations pour les aider à s'améliorer en évaluation formative, ou lorsque l'évaluation formative est appuyée par la technologie, les impacts sur les apprentissages et sur les rendements des élèves sont meilleurs (Andersson et Palm, 2017 ; Burns et coll., 2010 ; Kingston et Nash, 2011 ; Polly et coll., 2017 ; Vásquez et coll., 2017).

Les potentialités de l'évaluation formative établies dans les recherches quantitatives sont conformes aux résultats des études qualitatives et mixtes ayant examiné les perceptions et les expériences des enseignants et des élèves. Les différentes conclusions des études quantitatives, qualitatives et documentaires défient ainsi l'hypothèse développée par certains auteurs selon laquelle l'idée que l'utilisation de l'évaluation formative augmente directement les rendements scolaires est une sagesse conventionnelle (Dunn et Mulvenon, 2009).

Dans les recherches répertoriées dans la présente étude, l'efficacité de l'ensemble des pratiques de l'évaluation formative est étroitement liée aux conditions de leur mise en oeuvre. Les revues systématiques qui ont pris en compte les questions relatives aux conditions de mise en oeuvre de l'évaluation formative (Heitink et coll., 2016 ; Leong et coll., 2018) révèlent que certaines croyances et attitudes des enseignants et des élèves, ou encore, la pression des examens externes sont des obstacles qui limitent la mise en oeuvre efficace des pratiques liées à l'évaluation formative en classe.

Alors que ces conclusions sont partagées dans la méta-analyse de Black et Wiliam (1998), qui lie les effets positifs des évaluations formatives à leur qualité, les résultats des recherches quantitatives, qualitatives et mixtes (Burner, 2016 ; Büyükkarci, 2014 ; Noonan et Duncan, 2005 ; Saeed et coll., 2018) permettent de relever que, malgré la conscience de l'utilité des pratiques de l'évaluation formative, le manque de compétences des enseignants et l'immaturation des élèves du primaire et du secondaire sont discutés comme des défis qui limitent les pratiques effectives des activités liées à l'évaluation formative.

Le perfectionnement professionnel des enseignants, la formation des élèves, la modification des pratiques d'évaluation, le changement d'attitudes en rapport avec le rôle de l'élève dans l'enseignement-apprentissage sont autant de conditions pour implémenter les pratiques gagnantes de l'évaluation formative en salle de classe. À défaut, certains auteurs s'inquiètent de l'effet à long terme des pratiques d'évaluation formative mal comprises ou de la création d'écarts entre les élèves par l'évaluation formative qui n'est pas bien mise en oeuvre (Antonioni et James, 2014 ; Chouinard, Bowen, Cartier, Desbiens, Laurier, Plante et Butler, 2005).

7. Implications

Trois constats ressortent à la suite de cette revue systématique. Le premier constat est relatif aux contextes et aux disciplines concernées par les études. Dans l'ensemble, plus d'études ont été envisagées dans le contexte de l'école primaire où 14 études quantitatives, qualitatives et mixtes ont été réalisées contre sept études conduites dans le contexte du secondaire. Concernant les disciplines, l'évaluation formative a été envisagée dans plusieurs disciplines scolaires (mathématiques, lecture/écriture, français, biologie, anglais, informatique, éducation physique, sciences sociales). Les mathématiques et l'anglais au primaire se dégagent comme les disciplines où l'utilisation de l'évaluation formative est plus explorée dans les études. Il est important que des études soient envisagées dans les autres disciplines, surtout dans le contexte du secondaire.

Le deuxième constat distingue deux groupes de pratiques de l'évaluation formative à travers les recherches. L'autoévaluation, l'évaluation par les pairs, l'utilisation du portfolio, la carte conceptuelle, les grilles d'évaluation se révèlent comme des pratiques confirmées dans les recherches probantes. Des revues systématiques, des méta-analyses et des recherches quantitatives fournissent des preuves de leurs potentialités. Les résultats positifs des effets du développement des compétences d'autorégulation des élèves et des formats de la rétroaction (discussions interactives et orales) sur les apprentissages situent ces pratiques comme des pratiques émergentes et prometteuses en évaluation formative. Cependant, le nombre limité de recherches probantes les ayant examinées nécessite d'autres recherches expérimentales et longitudinales afin de les confirmer.

Le dernier constat est relatif au contexte géographique et à la langue des études analysées. Selon les zones géographiques et la langue, les pays européens, américains et asiatiques (enregistrant respectivement neuf, sept et six études) dominent les recherches probantes réalisées en évaluation des apprentissages, contrairement aux contextes africain et océanique. Des articles retenus, aucune étude n'est francophone. Cela implique que des efforts restent à faire dans ces contextes éducatifs.

8. Limites de la recherche

Cette recherche présente des limites qu'il convient de relever. La première limite de cette revue systématique est inhérente à toute recherche documentaire : l'exhaustivité. En effet, il se pourrait que cette recherche n'ait pas identifié toutes les études liées aux pratiques évaluatives probantes, d'autant plus que le critère linguistique (français et anglais seulement) excluait déjà des études potentiellement importantes menées dans d'autres langues. Ensuite, considérant qu'il existe plusieurs conceptualisations des composantes de l'évaluation formative, il a été impossible de définir un plan de concepts exhaustifs pour chercher le corpus des études dans les bases de données. Il se pourrait que, par le fait même, nos équations de recherche n'aient pas permis de couvrir totalement le champ conceptuel des pratiques évaluatives. Aussi l'approche utilisée n'a pas recouru à un accord d'interjuges pour évaluer les études analysées. Il est nécessaire de mener d'autres recherches documentaires sur les pratiques évaluatives probantes.

9. Conclusion

Cette recherche visait à documenter les pratiques évaluatives formatives probantes dans le contexte de l'enseignement primaire et secondaire. La recherche systématique a abouti à l'analyse de 35 articles révisés par les pairs. Les résultats révèlent que l'évaluation formative est un outil qui a un impact sur l'apprentissage des élèves, tant sur le processus que sur les résultats. Sept pratiques de l'évaluation formative, dont cinq confirmées et deux en développement, sont mises en évidence dans les études comme étant efficaces pour améliorer l'apprentissage des élèves : l'implication des apprenants dans le processus d'évaluation à travers les pratiques de l'autoévaluation, l'évaluation par les pairs, l'autorégulation, l'évaluation par le portfolio, l'utilisation des grilles d'évaluation, la carte conceptuelle et la rétroaction. Cependant, les études mettent leur efficacité en lien avec les conditions de mise en oeuvre. Les recherches montrent des défis en lien avec une mise efficace de ces pratiques d'évaluation dans les salles de classe. De nombreux facteurs de l'environnement scolaire, dont des facteurs liés aux compétences des enseignants et des élèves, la pression des examens et le manque de temps, limitent l'efficacité des pratiques de l'évaluation formative ou font en sorte que des tentatives de promotion de l'évaluation formative échouent dans certains contextes.

Il y a des implications pour les politiques, les pratiques et la recherche. Sur le plan des politiques et des pratiques, il est important que les politiques éducatives instaurent un environnement d'apprentissage favorable à l'émergence et à la pratique de l'évaluation formative par les enseignants. Au point de vue de la recherche, il est nécessaire de mener d'autres recherches documentaires systématiques sur les pratiques évaluatives probantes ainsi que des études longitudinales afin de confirmer l'efficacité des pratiques identifiées.



Noaga Dieudonné Kaboré
Étudiant à la maîtrise, Université Laval



Éric Frenette
Professeur, Université Laval



Marie-Hélène Hébert
Professeure, Université TÉLUQ

Note

[1] Dans cette présente recherche, l'usage épïcène du masculin a pour seul but d'alléger le texte.

Bibliographie

Les références précédées d'un astérisque font partie du corpus utilisé pour la recherche.

*Andersson, C. et Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development program. *Learning and instruction*, 49, 92-102. (<https://doi.org/10.1016/j.learninstruc.2016.12.006>) doi : <https://doi.org/10.1016/j.learninstruc.2016.12.006>

*Antoniou, P. et James, M. (2014). Exploring formative assessment in primary school classrooms: Developing a framework of actions and strategies. *Educational assessment, evaluation and accountability*, 26(2), 153-176. (<https://doi.org/10.1007/s11092-013-9188-4>) doi : <https://doi.org/10.1007/s11092-013-9188-4>

*Babinčáková, M., Ganajová, M., Sotáková, I. et Bernard, P. (2020). Influence of formative assessment classroom techniques (FACTs) on student's outcomes in chemistry at secondary school. *Journal of Baltic science education*, 19(1), 36-49. (<https://doi.org/10.33225/jbse/20.19.36>) doi : <https://doi.org/10.33225/jbse/20.19.36>

Bearman, M., Smith, C. D., Carbone, A., Slade, S., Baik, C., Hughes-Warrington, M. et Neumann, D. L. (2012). Systematic review methodology in higher education. *Higher education research & development*, 31(5), 625-640. (<https://doi.org/10.1080/07294360.2012.702735>) doi : <https://doi.org/10.1080/07294360.2012.702735>

Bissonnette, S., Gauthier, C. et Péladeau, N. (2010). Un objet qui manque à sa place : les données probantes dans l'enseignement et la formation. Dans B. Wentzel et M. Mellouki (dir.), *Actes de la recherche de la HEP-BEJUNE : n° 8. Recherche et formation à l'enseignement – Spécificités et interdépendance* (p. 107-133). Haute école pédagogique – BEJUNE.

*Black, P. et Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education: Principles, policy & practice*, 5(1), 7-74. (<https://doi.org/10.1080/0969595980050102>) doi : <https://doi.org/10.1080/0969595980050102>

*Boumediene, A. et Hamzaoui-Elachachi, H. (2017). The effects of formative assessment on Algerian secondary school pupils' text comprehension. *Arab world English journal*, 8(3), 172-190. (<https://dx.doi.org/10.24093/awej/vol8no3.12>) doi : <https://doi.org/10.24093/awej/vol8no3.12>

Broadfoot, P. et Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in education: Principles, policy & practice*, 11(1), 7-26. (<https://doi.org/10.1080/0969594042000208976>) doi : <https://doi.org/10.1080/0969594042000208976>

*Burner, T. (2014). The potential formative benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in educational evaluation*, 43, 139-149. (<https://doi.org/10.1016/j.stueduc.2014.03.002>) doi : <https://doi.org/10.1016/j.stueduc.2014.03.002>

*Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian journal of educational research*, 60(6), 626-648. (<https://doi.org/10.1080/00313831.2015.1066430>) doi : <https://doi.org/10.1080/00313831.2015.1066430>

*Burns, M. K., Klingbeil, D. A. et Ysseldyke, J. (2010). The effects of technology-enhanced formative evaluation on student performance on state accountability math tests. *Psychology in the schools*, 47(6), 582-591. (<https://doi.org/10.1002/pits.20492>) doi : <https://doi.org/10.1002/pits.20492>

*Butler, Y. G. et Lee J. (2010). The effects of self-assessment among young learners of English. *Language testing*, 27(1), 5-31. (<https://doi.org/10.1177/0265532209346370>) doi : <https://doi.org/10.1177/0265532209346370>

- *Büyükkarci, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International journal of instruction*, 7(1), 107-120.
- *Chang, C. C., Liang, C. et Chen, Y. H. (2013). Is learner self-assessment reliable and valid in a Web-based portfolio environment for high school students? *Computers & education*, 60(1), 325-334. (<https://doi.org/10.1016/j.compedu.2012.05.012>) doi : <https://doi.org/10.1016/j.compedu.2012.05.012>
- Chouinard, R., Bowen, F., Cartier, S. C., Desbiens, N., Laurier, M., Plante, I. et Butler, D. (2005). *L'effet de différentes approches évaluatives sur l'engagement et la persévérance scolaires dans le contexte du passage du primaire au secondaire. Rapport de recherche pour la subvention 2003-PS-94321*. Université de Montréal.
- Clanet, J. et Talbot, L. (2012). Analyse des pratiques d'enseignement. Éléments de cadrages théoriques et méthodologiques. *Phronesis*, 1(3), 4-18. (<https://doi.org/10.7202/1012560ar>) doi : <https://doi.org/10.7202/1012560ar>
- Dagenais, P. (dir.), Martin, V. et Renaud, J. (2013). *Les normes de production des revues systématiques. Guide méthodologique*. Institut national d'excellence en santé et en services sociaux (INESSS). (http://www.inesss.qc.ca/fileadmin/doc/INESSS/DocuMetho/INESSS_Normes_production_revues_systematiques.pdf)
- Deaudelin, C., Desjardins, J., Dezutter, O., Thomas, L., Corriveau, A., Lavoie, J., Bousadra, F. et Hébert, M. (2007). L'évaluation formative en contexte de renouveau pédagogique au primaire : analyse de pratiques au service de la réussite. *Nouveaux cahiers de la recherche en éducation*, 10(1), 27-45. (<https://doi.org/10.7202/1016856ar>) doi : <https://doi.org/10.7202/1016856ar>
- *Dunn, K. E. et Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical assessment, research, and evaluation*, 14(7), 1-11. (<https://doi.org/10.7275/jg4h-rb87>)
- Farza, L. (2018). La carte conceptuelle comme outil favorisant l'apprentissage de la modélisation des bases de données. *Revue internationale de pédagogie de l'enseignement supérieur*, 34(1), 1-21. (<https://doi.org/10.4000/ripes.1296>) doi : <https://doi.org/10.4000/ripes.1296>
- 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. (<https://doi.org/10.1080/02671520701296189>) doi : <https://doi.org/10.1080/02671520701296189>
- Grangeat, M. et Lepareur, C. (2019). Rôle du feedback enseignant sur l'autorégulation des apprentissages. *Évaluer : journal international de recherche en éducation et formation – e-JIREF*, 5(2), 5-28. (<http://hdl.handle.net/20.500.12162/3545>)
- Hattie, J. et Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112. (<https://doi.org/10.3102/003465430298487>) doi : <https://doi.org/10.3102/003465430298487>
- *Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K. et Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational research review*, 17, 50-62. (<https://doi.org/10.1016/j.edurev.2015.12.002>) doi : <https://doi.org/10.1016/j.edurev.2015.12.002>
- *Kingston, N. et Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational measurement: Issues and practice*, 30(4), 28-37. (<https://doi.org/10.1111/j.1745-3992.2011.00220.x>) doi : <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- *Kirton, A., Hallam, S., Peffers, J., Robertson, P. et Stobart, G. (2007). Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools. *British educational research journal*, 33(4), 605-627. (<https://www.jstor.org/stable/30032767>) doi : <https://doi.org/10.1080/01411920701434136>
- Landry, R., Becheikh, N., Amara, N., Ziam, S., Idrissi, O. et Castonguay, Y. (2008). *La recherche, comment s'y retrouver ? Revue systématique des écrits sur le transfert des connaissances en éducation*. Gouvernement du Québec. (http://www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/statistiques_info_decisionnelle/LaRechercheCommentSYRetrouver.pdf)

- Larose, F., Couturier, Y., Bédard, J. et 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(2), 31-48. (<https://doi.org/10.3917/lse.442.0031>) doi : <https://doi.org/10.3917/lse.442.0031>
- Laurier, M. D., Tousignant, R. et Morissette, D. (2005). *Les principes de la mesure et de l'évaluation des apprentissages* (3^e édition). Gaëtan Morin.
- Leclerc, K. (2013). *Analyse comparative des politiques d'évaluation des apprentissages canadiennes francophones : Manitoba, Nouveau-Brunswick, Ontario et Québec* [mémoire de maîtrise, Université Laval].
- *Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing writing*, 12(3), 180-198. (<https://doi.org/10.1016/j.asw.2008.02.003>) doi : <https://doi.org/10.1016/j.asw.2008.02.003>
- *Leong, W. S., Ismail, H., Costa, J. S. et Tan, H. B. (2018). Assessment for learning research in East Asian countries. *Studies in educational evaluation*, 59, 270-277. (<https://doi.org/10.1016/j.stueduc.2018.09.005>) doi : <https://doi.org/10.1016/j.stueduc.2018.09.005>
- *Mak, P. et Wong, K. (2018). Self-regulation through portfolio assessment in writing classrooms. *ELT journal*, 72(1), 49-61. (<https://doi.org/10.1093/elt/ccx012>) doi : <https://doi.org/10.1093/elt/ccx012>
- *McLaren, S. (2012). Assessment is for learning: Supporting feedback. *International journal of technology and design education*, 22, 227-245. (<https://doi.org/10.1007/s10798-011-9195-z>) doi : <https://doi.org/10.1007/s10798-011-9195-z>
- *McMillan, J. H., Venable, J. C. et Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical assessment, research, and evaluation*, 18(2), 1-15. (<https://doi.org/10.7275/tmwm-7792>)
- McQueen, D. V. et Anderson, L. M. (2000). Données probantes et évaluation des programmes en promotion de la santé. *Ruptures, revue transdisciplinaire en santé*, 7(1), 79-98.
- Ministère de l'Éducation du Québec (2003). *Politique d'évaluation des apprentissages : être évalué pour mieux apprendre, décroche tes rêves*. Gouvernement du Québec. (http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/evaluation/13-4602.pdf)
- Morissette, J. et Nadeau, M. (2011). Une typologie du savoir-faire enseignant en matière d'évaluation formative. *Mesure et évaluation en éducation*, 34 (3), 5-25. (<https://doi.org/10.7202/1024793ar>) doi : <https://doi.org/10.7202/1024793ar>
- Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages : enjeux pour l'enseignement*. De Boeck.
- Munn, Z., Stern, C., Aromataris, E., Lockwood, C. et Jordan, Z. (2018). What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC medical research methodology*, 18(5), 1-9. (<https://doi.org/10.1186/s12874-017-0468-4>) doi : <https://doi.org/10.1186/s12874-017-0468-4>
- *Ní Chróinín, D. et Cosgrave, C. (2013). Implementing formative assessment in primary physical education: Teacher perspectives and experiences. *Physical education and sport pedagogy*, 18(2), 219-233. (<http://dx.doi.org/10.1080/17408989.2012.666787>) doi : <https://doi.org/10.1080/17408989.2012.666787>
- *Noonan, B. et Duncan, C. R. (2005). Peer and self-assessment in high schools. *Practical assessment, research, and evaluation*, 10(17), 1-8. (<https://doi.org/10.7275/a166-vm41>)
- Organisation de coopération et de développement économiques (2005). *L'évaluation formative : pour un meilleur apprentissage dans les classes secondaires*. OCDE et CERI. (<https://doi.org/10.1787/9789264007420-fr>)
- *Panadero, E. et Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational research review*, 9, 129-144. (<https://doi.org/10.1016/j.edurev.2013.01.002>) doi : <https://doi.org/10.1016/j.edurev.2013.01.002>

- *Panadero, E., Tapia, J. A. et Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and individual differences*, 22(6), 806-813. (<https://doi.org/10.1016/j.lindif.2012.04.007>) doi : <https://doi.org/10.1016/j.lindif.2012.04.007>
- *Parr, J. M. et Timperley, H. S. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing writing*, 15(2), 68-85. (<https://doi.org/10.1016/j.asw.2010.05.004>) doi : <https://doi.org/10.1016/j.asw.2010.05.004>
- *Perels, F., Dignath, C. et Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European journal of psychology of education*, 24(1), 17-31. (<https://doi.org/10.1007/BF03173472>) doi : <https://doi.org/10.1007/BF03173472>
- *Phelan, J., Choi, K., Vendlinski, T., Baker, E. et Herman, J. (2011). Differential improvement in student understanding of mathematical principles following formative assessment intervention. *Journal of educational research*, 104(5), 330-339. (<https://doi.org/10.1080/00220671.2010.484030>) doi : <https://doi.org/10.1080/00220671.2010.484030>
- *Polly, D., Wang, C., Martin, C., Lambert, R. G., Pugalee, D. K. et Middleton, C. W. (2017). The influence of an Internet-based formative assessment tool on primary grades students' number sense achievement. *School science and mathematics*, 117(3/4), 127-136. (<https://doi.org/10.1111/ssm.12214>) doi : <https://doi.org/10.1111/ssm.12214>
- Rey, O. (2014). Entre laboratoire et terrain : comment la recherche fait ses preuves en éducation. *Dossier de veille de l'IFÉ*, (89), 1-28. (<https://edupass.hypotheses.org/13>)
- Rust, C., O'Donovan, B. et Price, M. (2005). A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment & evaluation in higher education*, 30(3), 231-240. (<https://doi.org/10.1080/02602930500063819>) doi : <https://doi.org/10.1080/02602930500063819>
- *Sadler, P. M. et Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational assessment*, 11(1), 1-31. doi : https://doi.org/10.1207/s15326977ea1101_1
- *Saeed, M., Tahir, H. et Latif, I. (2018). Teachers' perceptions about the use of classroom assessment techniques in elementary and secondary schools. *Bulletin of education and research*, 40(1), 115-130.
- *Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H. et Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of educational psychology*, 109(8), 1049-1066. (<https://psycnet.apa.org/doi/10.1037/edu0000190>) doi : <https://doi.org/10.1037/edu0000190>
- *Sebba, J., Crick, R. D., Yu, G., Lawson, H., Harlen, W. et Durant, K. (2008). *Systematic review of research evidence of the impact on students in secondary schools of self and peer assessment. Report no. 1614*. EPPi-Centre, Social Science Research Unit, Institute of Education, University of London.
- 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>) doi : <https://doi.org/10.1080/00461520.2019.1611432>
- *Stevenson, M. P., Hartmeyer, R. et Bentsen, P. (2017). Systematically reviewing the potential of concept mapping technologies to promote self-regulated learning in primary and secondary science education. *Educational research review*, 21, 1-16. (<https://doi.org/10.1016/j.edurev.2017.02.002>) doi : <https://doi.org/10.1016/j.edurev.2017.02.002>
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 83(10), 758-765. (<https://doi.org/10.1177/003172170208301010>) doi : <https://doi.org/10.1177/003172170208301010>
- *Vandergrift, L. (2000). Setting students up for success: Formative evaluation and FLES. *Foreign language annals*, 33(3), 290-304. (<https://doi.org/10.1111/j.1944-9720.2000.tb00607.x>) doi : <https://doi.org/10.1111/j.1944-9720.2000.tb00607.x>

*Vásquez, A., Nussbaum, M., Sciarresi, E., Martínez, T., Barahona, C. et Strasser, K. (2017). The impact of the technology used in formative assessment: The case of spelling. *Journal of educational computing research*, 54(8), 1142-1167. (<https://doi.org/10.1177/0735633116650971>) doi : <https://doi.org/10.1177/0735633116650971>

Wiliam, D. (2011). What is assessment for learning? *Studies in educational evaluation*, 37(1), 3-14. (<https://doi.org/10.1016/j.stueduc.2011.03.001>) doi : <https://doi.org/10.1016/j.stueduc.2011.03.001>

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, 45, 477-501. (<https://doi.org/10.1023/A:1023967026413>) doi : <https://doi.org/10.1023/A:1023967026413>

Résultats d'analyse détaillés incluant les articles retenus et leur méthodologie

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
1	Black, P. et Wiliam, D. (1998).	Analyser l'impact des pratiques de l'évaluation formative sur les apprentissages	Méta-analyse Synthèse de plus de 40 études	Discipline : / Royaume-Uni	L'attention est attirée sur la capacité de l'évaluation formative à jouer un rôle crucial dans l'amélioration d'un apprentissage efficace. L'impact de l'évaluation formative de la taille de l'effet est positif (effet de 0,4 à 0,7) sur l'apprentissage, peu importe le niveau scolaire. Les instruments jugés les plus efficaces : implication des élèves dans le processus d'évaluation à travers une rétroaction orale et écrite plutôt que des notes ou des appréciations, un partage des critères d'évaluation avec les élèves ; l'utilisation de l'autoévaluation et de l'évaluation par les pairs par les élèves.
2	Kingston, N. et Nash, B. (2011).	Déterminer si l'ampleur des effets signalés dans la documentation est conforme aux allégations couramment formulées au sujet de l'efficacité de l'évaluation formative	Méta-analyse Échantillon : 13 articles analysés (42 tailles d'effet indépendantes) Analyse : approche méta-analytique à effets aléatoires	Discipline : / École primaire, secondaire et supérieure États-Unis	Les effets de l'évaluation formative sont positifs : ampleur moyenne pondérée de l'effet : 0,20 ; médiane de l'ampleur de l'effet observé : 0,25. L'évaluation formative serait plus efficace en anglais qu'en mathématiques ou en sciences (ampleurs de l'effet : respectivement 0,32, 0,17, 0,09). L'évaluation formative fondée sur le perfectionnement professionnel et sur l'utilisation de systèmes informatisés est plus efficace que les autres approches.
3	Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H. et Cooper, H. (2017).	Examiner si et quand l'autoévaluation et l'évaluation par les pairs sont efficaces dans les écoles primaires et secondaires	Méta-analyse 33 articles analysés Analyse : analyse de contenu	Discipline : / École primaire et secondaire États-Unis	L'autoévaluation et l'évaluation par les pairs sont des pratiques d'évaluation formative efficaces qui aident les élèves du primaire et du secondaire à mieux réussir aux tests subséquents ; Les élèves ayant participé à l'autoévaluation ou à l'évaluation par les pairs obtiennent de meilleurs résultats aux tests subséquents que les élèves qui ne l'ont pas fait.
4	Dunn, K. E. et Mulvenon, S. W. (2009).	Proposer une analyse critique de la littérature fondamentale sur l'évaluation formative de Black et Wiliam (1998)	Examen critique de méta-analyse	Discipline : / États-Unis	Des problèmes méthodologiques et de population de certains articles dans les études de Black et Wiliam (1998) impliquent que les résultats de l'impact de l'évaluation formative sur le rendement général des élèves sont discutables. Nécessité de recherches plus empiriques dans le domaine de l'évaluation formative

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
5	McMillan, J. H., Venable, J. C. et Varier, D. (2013).	Analyser la qualité de la recherche incluse dans la méta-analyse de Kingston et Nash (2011)	Examen de critique de méta-analyse Une méta-analyse examinée Analyse : analyse de contenu	Discipline : / États-Unis	Des lacunes méthodologiques de plusieurs études incluses dans la méta-analyse sont relevées. Les conclusions de la méta-analyse concernant la relation entre l'évaluation formative et la réussite doivent être provisoires.
6	Burner, T. (2014).	Étudier le potentiel formatif de l'évaluation par le portfolio de l'élève dans le contexte de rédaction en langues secondes et en étrangère	Revue systématique 39 articles analysés Analyse de contenu	Discipline : / Norvège	Des moyens d'amélioration à la fois des processus et produits d'apprentissages : le portfolio facilite une rétroaction interactive ; il est moyen de révision des textes pour les élèves, l'évaluation est exempte d'anxiété. Il y a des impacts sur les rendements des élèves.
7	Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P. et Kippers, W. B. (2016).	Analyser les conditions préalables nécessaires à la mise en œuvre d'une pratique d'évaluation formative continue en classe	Revue systématique 25 articles analysés	Discipline : / Australie	Efficacité de l'évaluation formative, mais peu ou pas d'effets dans certaines études à cause des mises en œuvre inefficaces ; Des conditions liées aux connaissances et attitudes des enseignants et élèves évoquées comme préalables
8	Leong, W. S., Ismail, H., Costa, J. S. et Tan, H. B. (2018).	Examiner comment les chercheurs définissent l'évaluation pour l'apprentissage dans leurs études et les questions relatives à sa mise en œuvre dans le contexte des écoles primaires et secondaires en Asie de l'Est	Revue systématique 24 articles analysés Analyse : analyse de contenu	Discipline : / École primaire et secondaire Singapour	Même s'il existe différentes conceptualisations de l'évaluation formative, les chercheurs sont unanimes sur son but. On note des efforts en matière d'évaluation pour l'apprentissage dans les pays de l'Asie de l'Est, mais il existe de grands obstacles à sa mise en œuvre, dont la pression des examens externes, la culture indigène et les compétences des enseignants.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
9	Panadero, E. et Jonsson, A. (2013).	Examiner la recherche sur l'utilisation des grilles d'évaluation pour déterminer si et comment les grilles d'évaluation ont un impact sur l'apprentissage des élèves	Revue systématique Échantillon : 21 articles analysés Analyse : analyse de contenu	Discipline : / Espagne	Les grilles d'évaluation sont un moyen d'apprentissage et d'autorégulation soutenue (accroissement de la transparence de l'évaluation, aide au processus de rétroaction, amélioration de l'autonomie et de l'auto-efficacité des élèves). Les avantages des rubriques dépendent de leur format et de leur usage.
10	Sebba J., Crick R. D., Yu, G., Lawson H., Harlen, W. et Durant, K. (2008).	Examiner les données de recherche sur l'impact de l'autoévaluation et de l'évaluation par les pairs sur les élèves des écoles secondaires	Revue systématique 26 articles analysés Analyse : analyse de contenu	Discipline : / École secondaire Royaume-Uni	Des résultats positifs de l'autoévaluation et de l'évaluation par les pairs sur le plan du rendement scolaire, de l'estime de soi, de leur engagement accru envers l'apprentissage. Il n'y a pas de résultats clairs entre l'appropriation du processus par les élèves et les résultats positifs ; des différences significatives entre les différents groupes d'élèves selon le sexe, l'origine ethnique ou les résultats antérieurs
11	Stevenson, M. P., Hartmeyer, R. et Bentsen, P. (2017).	Examiner comment les technologies de cartographie conceptuelle influencent l'apprentissage autorégulé	Revue systématique 17 articles analysés Analyse : analyse de contenu	Discipline : / Danemark	Les technologies de cartographie conceptuelle peuvent influencer positivement sur l'apprentissage autorégulé en améliorant les compétences cognitives, métacognitives et de motivation à divers degrés. Les logiciels semblent être particulièrement utiles pour développer des compétences cognitives par leur facilité d'utilisation.
12	Andersson, C. et Palm, T. (2017).	Examiner les effets d'un programme de développement professionnel des enseignants de mathématiques en évaluation formative	Quantitative Échantillon : enseignants (n = 45) ; élèves : n = 738 (4 ^e année primaire) Collecte de données : questionnaire Analyse : tests ANCOVA	Discipline : mathématiques École primaire Suède	Les pratiques d'évaluation formative des enseignants en situation de classe se sont améliorées. Les impacts de l'amélioration sont positifs sur différents processus mathématiques et sur les rendements des élèves.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
13	Babincáková, M., Ganajová, M., Šotáková, I. et Bernard, P. (2020)	Évaluer l'impact de l'utilisation pendant dix leçons de chimie de l'évaluation formative axée sur l'autoévaluation et l'évaluation par les pairs sur les résultats des élèves dans un test sommatif.	Quantitative Échantillon aléatoire : élèves (n = 202 ; groupe expérimental, n = 105 ; groupe contrôle, n = 9) Écoles, n = 4 ; classes, n = 3 Collecte de données : questionnaires et tests Analyse : test du chi carré	Discipline : chimie École secondaire Slovaquie	Différence des résultats obtenus par le groupe expérimental et le groupe contrôle statistiquement significative. L'évaluation formative stimule les compétences cognitives d'ordre supérieur sur le plan de l'application et de l'analyse. Les élèves considèrent que l'évaluation formative utile pour trouver leurs lacunes et améliorer leurs résultats scolaires.
14	Boumediene, A. et Hamzaoui-Elachachi, H. (2017).	Étudier l'impact de l'évaluation formative sur la compréhension du texte des élèves ainsi que sur leurs résultats et leurs attitudes	Quantitative Échantillon, élèves (n = 32) Collecte de données : questionnaires et tests Analyse : Statistiques descriptives	Discipline : français École secondaire Algérie	Des progrès significatifs des résultats des élèves (différence de scores moyens de 2,98 entre au post-test) sont enregistrés sur quatre compétences linguistiques : compétences grammaticales, textuelles, fonctionnelles et sociolinguistiques ; La rétroaction identifiée comme sources des gains. Les élèves jugent l'intervention formative importante pour leur compréhension des textes.
15	Burns, M. K., Klingbeil, D. A. et Yaseldyke, J. (2010).	Examiner les effets de l'utilisation d'un programme d'évaluation formative assistée par technologie sur le pourcentage d'élèves ayant obtenu des résultats satisfaisants aux évaluations sommatives des mathématiques	Quantitative Échantillon : 360 écoles. Collecte de données : données statistiques des résultats aux examens Analyse : tests ANCOVA	Discipline : lecture et mathématiques École primaire États-Unis	Augmentation du niveau de rendements des élèves grâce à l'évaluation formative assistée par la technologie ; il y a un effet positif sur le pourcentage d'élèves qui réussissent aux évaluations sommatives.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
16	Chang, C. C., Liang, C. et Chen, Y. H. (2013).	Explorer la fiabilité et la validité de l'autoévaluation des portefeuilles sur le Web	Quantitative Échantillon : élèves, n = 72 Collecte de données : tests Analyse : analyses de corrélation et de test t	Discipline : cours d'application informatique École secondaire Taiwan	L'autoévaluation est une méthode dont la fiabilité et la validité sont appropriées. Il y a une cohérence importante entre la notation des élèves et celle des enseignants dans des aspects distincts.
17	Panadero, E., Tapia, J. A. et Huertas, J. A. (2012).	Comparer les effets de deux outils d'autoévaluation différents (les rubriques et les scripts) sur l'autorégulation, l'apprentissage et l'auto-efficacité	Quantitative Échantillon : élèves, n = 120 Collecte de données : questionnaires Analyse : ANOVA à sens unique, puis à mesures répétées	Discipline : / École secondaire Espagne	Effets positifs des deux outils : l'utilisation des outils d'autoévaluation favorise un niveau d'autorégulation plus élevé que l'absence d'autoévaluation et l'utilisation des deux outils d'autoévaluation augmente l'apprentissage au sein du groupe témoin. Les textes améliorent davantage l'autorégulation que les rubriques.
18	Perels, F., Dignath, C. et Schmitz, B. (2009).	Examiner les effets de la formation des compétences d'autorégulation des élèves dans les classes de mathématiques sur l'apprentissage autorégulé ainsi que sur la réussite en mathématiques	Quantitative Échantillon : enseignante, n = 1 Élèves, n = 55 Collecte de données : questionnaires Analyse : analyse de variance, avec le temps comme facteur de mesure répétée	Discipline : mathématiques École primaire Allemagne	Les compétences et les rendements sont soutenus par l'autorégulation : les deux classes ont amélioré leurs compétences en mathématiques sur les diviseurs et les multiplicateurs après l'intervention.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
19	Phefan, J., Choi, K., Vendinski, T., Baker, E. et Herman, J. (2011).	Démontrer l'amélioration différentielle de la compréhension des principes mathématiques par les élèves à la suite d'une intervention d'évaluation formative à l'école primaire	Quantitative Échantillon : enseignants, n = 85 ; Élèves : n = 4 091 ; écoles : n = 27 Collecte de données : tests Analyse : statistiques descriptives et Modèle statistique ; Modèle hiérarchique à deux niveaux	Discipline : mathématiques École primaire États-Unis	Les résultats des élèves en traitement dans les conceptions intra- et interécoles sont nettement supérieurs à ceux des élèves des groupes témoins, et les effets sont plus importants pour ceux aux résultats du prétest supérieurs.
20	Polly, D., Wang, C., Martin, C., Lambert, R. G., Pugalee, D. K. et Middleton, C. W. (2017).	Examiner la relation entre l'utilisation par les enseignants d'évaluation formative basés sur internet et le rendement des élèves en mathématiques dans le cadre d'un projet de perfectionnement professionnel	Quantitative Échantillon : enseignants, n = 906 ; élèves, n = 13 567 élèves de la maternelle à la 2 ^e année du primaire (6 % en maternelle, 43 % en première année et 51 % en deuxième année). Collectes de données : tests Analyse : modèles linéaires hiérarchiques	Discipline : mathématiques De la maternelle à la 2 ^e année du primaire États-Unis	Tous les élèves du groupe expérimental ont connu une croissance. Le rendement des élèves était influencé par la fréquence de l'évaluation, le statut socioéconomique et la taille des écoles. Les élèves qui ont reçu plus d'évaluation, ainsi que ceux issus de milieux défavorisés et de grands établissements ont obtenu plus de gains que leurs pairs.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
21	Sadler, P.M. et Good, E. (2006).	Examiner la validité et la fiabilité de l'autoévaluation et de l'évaluation par les pairs et leurs impacts sur l'apprentissage des élèves en classes de sciences générales	Quantitative Échantillon : élèves : n = 101 (classes : n = 4) Collectes de données : tests Analyse : statistiques descriptives et modèles d'analyse de la variance	Discipline : sciences générales (biologie) École secondaire États-Unis	Il y a de fortes corrélations entre les notes attribuées par les élèves et celles attribuées par l'enseignant suggérant que les notes des élèves sont valides et fiables. Des gains significatifs et importants observés pour la classe qui s'évaluent seulement. L'évaluation par les pairs n'a pas donné lieu à des gains importants.
22	Saeed, M., Tahir, H. et Latif, I. (2018).	Reconnaitre les principales méthodes utilisées par le personnel enseignant pour évaluer l'apprentissage des élèves au primaire et au secondaire dans le district de Lahore	Quantitative Échantillon : enseignants, n = 500 Collectes de données : questionnaire Analyse : statistiques descriptives	Discipline : / École primaire et secondaire Pakistan	Dans la plupart des écoles publiques et privées, les enseignants utilisent l'évaluation sommative. Ils pensent que l'évaluation formative et sommative peut jouer un rôle plus crucial dans la promotion de l'apprentissage des élèves en classe.
23	Vásquez, A., Nussbaum, M., Sciarresi, E., Martínez, T., Barahona, C. et Strasser, K. (2017).	Examiner l'impact de différentes technologies sur l'apprentissage des élèves lorsqu'on utilise l'évaluation formative pour enseigner l'orthographe	Quantitative Échantillon : élèves, n = 90 Collecte de données : résultats à un test écrit au prétest et post-test ; rapport de progrès de l'élève généré par logiciel sur des tablettes et ordinateurs interpersonnels. Analyse : analyses corrélationnelles et tests ANOVA et ANCOVA	Discipline : orthographe École primaire Chili	Une stratégie d'apprentissage assistée par la technologie, utilisant l'évaluation formative adaptée aux besoins de l'élève produit plus de gains significatifs d'apprentissage qu'une stratégie d'examen traditionnel établie par l'autorité et ne prenant pas en compte les besoins particuliers des élèves. Des améliorations de l'apprentissage ont été observées chez les élèves utilisant des tablettes que chez ceux utilisant l'ordinateur interpersonnel.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
24	Antoniou, P. et James, M. (2014).	Explorer la perception et la mise en œuvre de l'évaluation formative dans les classes du primaire	Qualitative Échantillon : enseignants (n = 4) ; élèves : n = 21 à 25 dans chacune des classes (n = 4) Collecte de données : entretiens, observations en classe, une analyse documentaire Analyse : comparaison constante	Discipline : diverses matières du programme École primaire Chypre	Les enseignants ont des perceptions positives des valeurs de l'utilisation des pratiques d'évaluation formative, mais leur vision de l'évaluation formative est incomplète.
25	Lee, I. (2007).	Examiner la nature et les fonctions de la rétroaction des enseignants afin de déterminer dans quelle mesure cette rétroaction sert à l'évaluation à des fins d'apprentissage	Qualitative Échantillon : professeurs d'anglais : n = 26 ; élèves : n = 174 Collecte de données : textes d'élèves ; entretiens Analyse : analyse de contenu et analyse par thèmes émergents	Discipline : / École secondaire Hong Kong	La rétroaction a eu un effet positif ou négatif suivant sa nature, sa qualité et sa variété : les élèves souhaitent avoir des rétroactions orales et des discussions interactives.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
26	Mak, P., Wong K. (2018).	Examiner comment les enseignants du primaire encouragent l'autorégulation des élèves par le biais de l'évaluation par le portfolio	Qualitative Échantillon : enseignants, n = 2 ; élèves, n = 69 Collecte de données : entretiens semi-structurés, observations en classe, analyse documentaire du matériel pédagogique Analyse : Analyse de contenu	Discipline : / École primaire Hong Kong	L'usage des portfolios par les étudiants contribue au développement d'un apprentissage autorégulé. Les élèves pensent être devenus des apprenants indépendants capables d'évaluer et de surveiller leur propre travail à travers la pratique de l'autoévaluation, de l'évaluation par les pairs et les rétroactions.
27	McLaren, S. (2012).	Examiner les opportunités offertes par une approche <i>e-scape</i> (<i>e-solutions for creative assessment in portfolio environment</i>) en vue de soutenir l'évaluation de l'apprentissage	Qualitative Échantillon : enseignants, n = 5 Élèves au secondaire : n = 305 Collecte de données : forums électroniques et des entretiens ; Analyse : analyse de contenu	Discipline : / École primaire et secondaire Écosse	Les portfolios électroniques fournissent des preuves utiles pour le diagnostic et l'évaluation formative.
28	Ni Chróinín, D. et Cosgrave, C. (2013).	Explorer les expériences d'utilisation de l'évaluation formative par des enseignants d'éducation physique dans leurs classes	Qualitative Échantillon : enseignants, n = 5 Collecte de données : journaux réfléchissants ; focus group Analyse : comparaison constante	Discipline : éducation physique École primaire Irlande	L'utilisation de variété de pratiques évaluatives formatives (évaluation par les pairs, les critères d'évaluation, rétroaction) dans les cours d'éducation physique a permis de structurer les processus de planification, d'enseignement et d'apprentissage. Cela a eu un impact positif sur les apprentissages.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
29	Noonan, B. et Duncan, C. R. (2005).	Explorer la nature et la fréquence de l'utilisation de l'autoévaluation et de l'évaluation par les pairs comme pratiques d'évaluation en classe	Qualitative Échantillon : enseignants, n = 118 Collecte de données : questionnaires Analyse : comparaison constante	Discipline : mathématiques, sciences, anglais, études sociales École secondaire Canada	De nombreux enseignants utilisent l'évaluation par les pairs et l'autoévaluation en classe. Dans toutes les matières, les enseignants les perçoivent positivement et pensent qu'il faut améliorer leur applicabilité en classe. Selon certains enseignants, l'une des raisons de ne pas utiliser l'évaluation par les pairs et l'autoévaluation est l'imaturité des élèves du secondaire pour son usage.
30	Burner, T. (2016).	Examiner comment les enseignants et les élèves perçoivent et expérimentent l'évaluation formative en anglais langue étrangère	Approche mixte Échantillon : enseignants, n = 4 ; Élèves, n = 100 Collecte de données : questionnaires et entretiens Analyse : statistiques descriptives et analyse de thèmes émergents	Discipline : Anglais langue étrangère. École intermédiaire Norvège	Les opinions des enseignants et celles des élèves sur la pratique de l'évaluation formative de l'écriture au sujet des commentaires, des notes, de la révision du texte, de l'autoévaluation et des participations sont contradictoires. Les élèves apprécient les rétroactions constructives, orales et interactives. L'autoévaluation est difficile pour les élèves les plus jeunes (huitième année).
31	Butler, Y. G. et Lee J. (2010).	Examiner l'efficacité de l'autoévaluation dans un programme d'anglais langue étrangère	Approche mixte Échantillon : élèves, n = 254 Classes : n = 8 (4 pour chaque école) Collecte de données : tests et entretiens Analyse : statistiques descriptives, tests ANOVA	Discipline : anglais langue étrangère École primaire Corée du Sud	Capacités accrues d'autoévaluation des élèves ; Des effets positifs sur les performances en anglais ainsi que sur leur confiance dans l'apprentissage. Enseignants et élèves perçoivent différemment l'efficacité de l'autoévaluation en fonction du contexte d'enseignement et d'apprentissage.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
32	Büyükkarci, K. (2014).	Montrer les croyances des enseignants d'anglais du primaire sur l'évaluation formative	Approche mixte Échantillon : enseignants d'anglais du primaire (n = 69) sélectionnés Collecte de données : questionnaires et entretiens Analyse : statistiques descriptives, tests T	Discipline : Anglais École primaire Turquie	Les professeurs d'anglais ont une bonne perception à l'égard de l'évaluation formative. La rétroaction, le partage des objectifs d'apprentissage, l'évaluation par les pairs et l'autoévaluation jugés utiles au processus d'apprentissage. Les enseignants ne peuvent pas les utiliser fréquemment dans chaque classe.
33	Kirton, A., Hallam, S., Peffers, J., Robertson, P. et Stobart, G. (2007).	Examiner la mise en œuvre dans différentes écoles, ainsi que les effets d'un projet de soutien à la pratique professionnelle en évaluation formative	Approche mixte Échantillon : 35 écoles Collecte de données : preuves documentaires, entretiens. Analyse : statistiques descriptives et analyse de contenu	Discipline : / École primaire et secondaire Royaume-Uni	Des preuves particulières de l'amélioration du travail des élèves, de leur motivation, de leur confiance en soi, de leur participation, de leurs résultats scolaires. La majorité des écoles participantes ont adopté sans réserve les pratiques d'évaluation formative.
34	Parr, J. M. et Timperley, H. S. (2010).	Définir les caractéristiques de la rétroaction de qualité Présenter des données probantes démontrant que la capacité d'un enseignant à donner cette rétroaction écrite de qualité a un impact sur les progrès des élèves en écriture	Approche mixte Échantillon : enseignants, n = 59 ; élèves, n = 375. Collecte de données : questionnaire, observation directe, données sur le rendement de l'écriture Analyse : test t, analyse de contenu	Discipline : lecture/écriture École primaire Nouvelle-Zélande	La rétroaction formative de qualité développe l'écriture des élèves : relation significative entre le score de qualité de la rétroaction et l'étendue moyenne des gains en écriture des élèves de leur classe ($r = 0,685, p < 0,01$). Les élèves ont jugé la rétroaction écrite moins utile par rapport aux discussions interactives.

N°	Articles	Objectifs	Méthodologie	Contexte	Résultats
35	Vandergrift, L. (2000).	Vérifier la capacité des instruments d'évaluation d'impacter positivement les résultats des élèves dans le cadre du NCFAP (National Core French Assessment Project)	Approche mixte Échantillon : enseignants, n = 162 ; Élèves, n = 148 Collectes de données : questionnaires (enseignants) commentaires (élèves) Analyse : analyse de contenu	Discipline : français École primaire Canada	Sur 33 instruments expérimentés, ceux identifiés par les enseignants et les élèves comme particulièrement flexibles, adaptés et ayant un potentiel impact positif sur leur responsabilisation, leur motivation et leur autonomie sont le retour réflexif (des élèves), l'autoévaluation et l'évaluation par les pairs, les listes de vérification, les grilles d'observation, les échelles de notation et la complétion de dossier.