

Comment aider les élèves en difficulté en mathématiques ?

Les réponses de Campbell et du What Works Clearinghouse (WWC)

Nathalie ROQUES

Aout 2021

Résumé

La sélection, le codage et les calculs des tailles d'effet des études primaires concernées par deux méta-analyses réalisées par Campbell et par le What Works Clearinghouse (WWC) ont été comparés. Publiées toutes les deux au début de l'année 2021, ces synthèses posent une même question : définir les interventions efficaces en direction des élèves en difficulté en mathématiques et scolarisés en primaire. Dix-neuf études communes ont été recensées. Presque toutes sont caractérisées comme associées à un enseignement en petits groupes par Campbell et à un enseignement explicite par le WWC. L'apprentissage soutenu par les pairs présenté par Campbell comme ayant un impact global positif important semble peu concerner les mathématiques dans cette synthèse, et ne figure pas parmi les recommandations proposées par le WWC.

Avant-propos

Le terme *méta-analyse* utilisé dans le texte fait référence à la synthèse quantitative d'études primaires sélectionnées selon des critères précisément définis (COOPER, 2019). Pour alléger le texte, le terme *synthèse* sera parfois employé comme synonyme.

Ce document a été écrit après la rédaction de *Comment enseigner les maths ? La réponse du What Works Clearinghouse* qui fournit des informations sur les principes suivis par les méta-analyses et qui rend compte également de quelques questions et discussions relatives à ce type de synthèses. Ces éléments ne seront pas repris ici.

Certaines traductions de l'américain sont suivies de l'expression originale.

Des documents annexes (fichiers Excel) peuvent être téléchargés sur le site www.mathadoc.fr. A la fin du texte, un glossaire donne les définitions des mots écrits en **gras** à leur première apparition dans le texte.

Introduction

La **sélection**, le **codage** et les calculs des **tailles d'effet** des **études primaires** concernées par deux **méta-analyses a priori** assez proches, ont été comparés. Publiées toutes les deux au début de l'année 2021, ces synthèses posent une même question : définir les **interventions** efficaces en direction des élèves en difficulté en mathématiques et scolarisés en primaire. Dans les deux cas, les interventions se déroulent dans les établissements scolaires. La première méta-analyse intitulée *Targeted school-based interventions for improving reading and mathematics for students with or at risk of academic difficulties in Grades K-6 : a systematic review* (DIETRICHSON, 2021), a été publiée par [Campbell](#), un réseau international qui produit des synthèses dans de nombreux domaines des sciences sociales ; le rapport publié adopte la structure classique d'un article scientifique (avec notamment des parties méthodes, résultats et discussion clairement identifiées) à l'attention d'un public de spécialistes. Cette méta-analyse s'est intéressée aux interventions portant sur les mathématiques mais aussi sur la lecture. La seconde méta-analyse intitulée *Assisting students struggling with mathematics : intervention in the elementary grades* (FUCHS, 2021), a été publiée par le [What Works Clearinghouse](#) (WWC), une émanation de l'Institut des sciences de l'éducation du département de l'Éducation des États-Unis, sous la forme d'un Guide des Pratiques à destination des enseignants. Cette synthèse s'organise autour de 6 recommandations et ne concerne cette fois que l'enseignement des mathématiques ; des annexes présentent les résultats statistiques et les méthodes suivies qui ont permis aux auteurs de conclure. Pour ces deux méta-analyse, un protocole a été défini et publié séparément.

On s'est intéressé ici plus particulièrement à la sélection, au codage puis aux calculs des tailles d'effet des études primaires, c'est-à-dire aux premières étapes d'une méta-analyse, dans l'objectif de comparer les méthodes mises en œuvre dans ces deux synthèses (pour une description générale des objectifs et procédures des méta-analyse, voir www.mathadoc.fr).

Dans un second temps, l'absence d'une caractéristique d'enseignement (l'apprentissage soutenu par les airs) des études sélectionnées communes à ces deux méta-analyses a été interrogée.

Méthodes

Les informations fournies par les deux méta-analyses ont été analysées et ont permis de caractériser les études sélectionnées, les interventions ou les tailles d'effet calculées en fonction de paramètres choisis par l'auteur de ce texte. Certains des résultats de ces deux synthèses (des pourcentages) ont été recalculés ; d'autres ont été calculés ici pour la première fois (comme certaines valeurs-p des tailles d'effet).

Des fichiers Excel sont téléchargeables sur www.mathadoc.fr qui détaillent tous ces calculs.

Les études sélectionnées et codées par Campbell

La sélection

La première étape ici consiste à rechercher, à partir de critères définis *a priori* et explicités dans le protocole, les études primaires **éligibles**. Dans la méta-analyse de Campbell, les études éligibles sont soit des essais contrôlés randomisés, soit des études quasi-expérimentales. Elles doivent avoir été publiées entre janvier 1980 et juin 2018 et rédigées en anglais, allemand, suédois, norvégien ou danois. Les interventions doivent concerner des élèves repérés comme ayant des difficultés dans les apprentissages académiques. Les tests utilisés doivent être standardisés ce qui explique que les résultats issus de tests développés par les chercheurs (tests « maison ») n'ont pas été inclus dans la méta-analyse. La **validité interne** des études éligibles est ensuite évaluée, et les études dont la qualité méthodologique est jugée insuffisante ne sont pas sélectionnées. Les études non sélectionnées sont listées et des motifs d'exclusion sont proposés (DIETRICHSON, 2021, p.56)¹.

Le codage

Chaque étude² sélectionnée a analysé une ou plusieurs interventions (encore dénommé traitement). Dans ce dernier cas, des échantillons d'élèves différents ont chacun subi un traitement différent. Chaque intervention (ou traitement) aboutie au calcul d'une ou de plusieurs tailles d'effet. Dans ce dernier cas, des tests différents³ ont pu être utilisés pour évaluer les élèves ayant subi un même traitement, ou bien un même test a été mis en œuvre à des moments différents (tout de suite après l'intervention ou plus longtemps après par exemple). Les interventions ont la plupart du temps ciblé des **domaines**, et parmi eux 7 concernent les mathématiques. Elles ont également été associées à des éléments pédagogiques caractéristiques. Une intervention peut être associée à plusieurs **caractéristiques pédagogiques** et/ou à plusieurs domaines mathématiques (**schéma 1**). Les études associées à une même caractéristique pédagogique ou à un même domaine constituent alors un groupe pour lequel une taille d'effet globale a été calculée⁴. Dans l'exemple fictif du **schéma 1**, la caractéristique 1 a été analysée lors de deux interventions (1 et 3) mises en œuvre par deux études (A et B), et a donné lieu au calcul de 3 tailles d'effet (ES 1, ES 2 et ES 6)⁵ qui permettront d'évaluer l'effet global de cette caractéristique. Le domaine 2 a été ciblé par deux interventions (1 et 2) mise en œuvre par une étude (A) et a donné lieu au calcul de 5 tailles d'effet (ES 1, ES 2, ES 3, ES 4, ES 5) qui permettront d'évaluer l'effet global des interventions dans ce domaine.

¹ La liste des études non éligibles (trop longue) n'est pas publiée.

² La plupart du temps, une étude est équivalente à un document publié (article dans une revue, thèse, ...). Il arrive que plusieurs articles concernent une même étude et c'est pour cette raison que Campbell parle de clusters d'études.

³ Les tests ont pu concerner soit les mathématiques soit la lecture.

⁴ C'est le modèle des effets aléatoires qui est utilisé (voir www.mathadoc.fr)

⁵ ES = tailles d'effet (*Effect Size*)

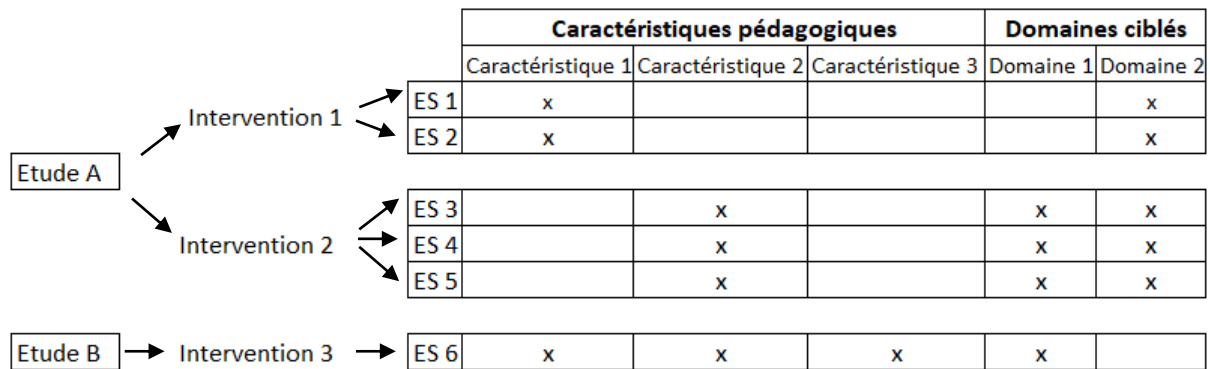


Schéma 1 : éléments descriptifs de la méta-analyse de Campbell

Les informations concernant les études sélectionnées sont répertoriées dans un fichier csv¹. Chaque ligne contient les données associées à une taille d'effet calculée. Certaines concernent l'étude (les noms des auteurs, la langue, le design par exemple), d'autres concernent l'intervention (les caractéristiques pédagogiques et les domaines ciblés) et d'autres encore la taille d'effet calculée (la taille des échantillons, la nature du test -lecture et/ou mathématiques). Nous nous intéresserons plus particulièrement ici aux 8 caractéristiques pédagogiques et aux 7 domaines mathématiques décrits ci-dessous et qui sont donc associés aux interventions.

Caractéristiques pédagogiques

Version originale	Traduction
<i>Computer Assisted Instruction (CAI)</i>	Enseignement assisté par ordinateur (utilisation d'un ordinateur, tablette, ...)
<i>Coaching of personnel</i>	Formation professionnelle ciblée sur l'implémentation d'un programme spécifique à la lecture ou aux mathématiques en direction d'élèves en difficulté.
<i>Incentives</i>	Récompense des élèves (financière ou autre)
<i>Medium group instruction</i>	Élèves travaillant en groupes de taille moyenne (entre 6 et 20 élèves)
<i>Other method</i>	Autres méthodes
<i>Peer assisted instruction</i>	Apprentissage soutenu par des pairs (les pairs sont des élèves pouvant être scolarisés jusqu'en terminale).
<i>Progress monitoring</i>	Suivi des progrès
<i>Small group instruction</i>	Élèves travaillant en groupes de petite taille (entre 1 et 5 élèves)

¹ Ce fichier est l'annexe h, téléchargeable sur le site internet (onglet *Supportive information* en bas de l'article).

Domaines mathématiques

Version originale	Traduction
<i>Algebra, PreAlgebra</i>	Algèbre, initiation aux équations, graphiques
<i>Fractions</i>	Fractions
<i>Geometry</i>	Géométrie (forme, taille, position, ...)
<i>Multiple math</i>	Plus d'un domaine mathématique est concerné, ou bien l'intervention porte sur des notions mathématiques plus générales
<i>Number sense</i>	Connaissance des nombres : comptage, reconnaissance de nombres, comparaison
<i>Operations</i>	Compétences calculatoires (addition, soustraction, multiplication, division)
<i>Problem solving</i>	Résolution de problèmes : trouver les solutions à des problèmes verbaux

Résultats généraux

Les principaux résultats publiés par Campbell concernent à la fois les mathématiques et la lecture. La distinction entre ces deux disciplines a été jugée peu pertinente par les auteurs et très peu de résultats concernent spécifiquement les mathématiques. Deux caractéristiques pédagogiques sont mises en avant : l'apprentissage soutenu par les pairs (*peer-assisted instruction*) avec une taille d'effet globale égale à 0,44 (CI= [0,28 ; 0,61])¹ et l'apprentissage en petits groupes (*small group instruction*) avec une taille d'effet globale égale à 0,38 (CI = [0,31 ; 0,44]). Ce sont 1 334 tailles d'effets qui ont été calculées à partir des données de 327 interventions publiées dans 195 études sélectionnées. Les noms des interventions ne sont pas mentionnés dans le fichier csv publié, mais l'auteur m'a transmis une version comprenant cette information. C'est une donnée importante, car ce sont les interventions (et non les tailles d'effet ou les études) qui sont associées à une caractéristique pédagogique ou à un domaine, et les premiers résultats présentés dans le tableau 1 de l'article publié (DIETRICHSON, 2021, p.17) se basent sur les nombres d'interventions comprises dans chaque catégorie. Si j'ai bien dénombré 1334 tailles d'effets et 195 études, j'ai par contre compté 294 interventions² (ou traitement) et non 327 comme indiqué dans l'article.

Les études sélectionnées en maths

A partir du fichier csv, j'ai dénombré 239 tailles d'effet, 103 interventions et 67 études associées à un test en mathématiques. Parmi elles, 209 tailles d'effets, 90 interventions et 58 études sont également associées à au moins un des 7 domaines mathématiques³ ; cela veut donc dire que 30 tailles d'effets, 13 interventions et 9 études ne sont pas associées à un domaine mathématique. Dans ce cas, l'intervention est associée à un domaine plus général ou encore au développement des compétences métacognitives. A l'inverse, quand on sélectionne les tailles d'effet ayant comme cible au moins un domaine mathématique, 7 sont

¹ CI = *Confidence Interval* (intervalle de confiance à 95 %).

² En supprimant les doublons de la liste des interventions.

³ J'ai rajouté une colonne AZ au fichier csv afin de les identifier (voir annexe 4 sur www.mathadoc.fr).

associées à un test en lecture et non à un test en mathématiques : cela vient du fait que certaines interventions ont concerné les mathématiques *et* la lecture (**schéma 2**).

On constate ici que, en considérant les tailles d'effet ou les interventions ou les études, les mathématiques sont moins souvent ciblées par les chercheurs que la lecture (voir **tableau 1**). Par exemple 18 % des tailles d'effet calculées sont associées à un test en mathématiques¹.

	Tailles d'effets	Interventions	Études
Nombre total	1334	327 (ou 294)	195
Test = maths (nombre)	239	103	67
Test = maths (%)	18 %	31 % (ou 35 %)	34 %
Test et domaine = maths (nombre)	209	90	58
Test et domaine = maths (%)	16 %	28 % (ou 31 %)	30 %

Tableau 1 : nombre de tailles d'effets, d'interventions et d'études en mathématiques

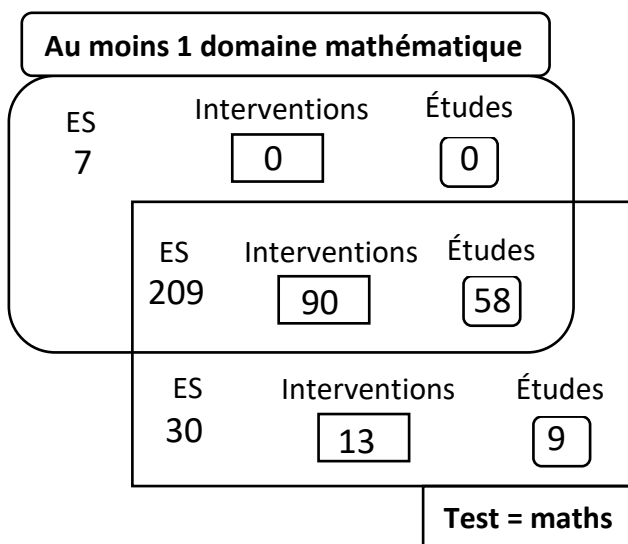


Schéma 2 : nombre de tailles d'effet, d'interventions et d'études concernées par les mathématiques

¹ C'est une valeur recalculée ici et également celle qui est mentionnée dans le texte de l'article (DIETRICHSON, 2021, p. 27).

Les études sélectionnées et codées par le What Works Clearinghouse (WWC)

La sélection

Les études éligibles sont soit des essais contrôlés randomisés, soit des études quasi-expérimentales. Elles doivent avoir été publiées entre janvier 2004 et décembre 2019¹ et rédigées en anglais. Les interventions doivent concerner des élèves en difficulté en mathématiques. Les critères d'éligibilité spécifiques à ce guide des pratiques sont détaillés dans un protocole. Seules les interventions sur des groupes d'élèves (soit en individuel, soit en petits groupes de 2 à 6 élèves) ont été retenues et donc les interventions en classe entière n'ont pas été sélectionnées. Les tests peuvent être des tests nationaux standardisés mais aussi des tests développés par les chercheurs (tests « maison »). La validité interne des études éligibles est ensuite évaluée en fonction de critères définis dans le WWC Standards Handbook (4.0). Elles peuvent être soit conformes sans restriction aux normes WWC, soit conforme avec restriction aux normes WWC, soit non conforme aux normes WWC. Dans ce dernier cas, l'étude est exclue de la méta-analyse. La liste des études non éligibles comme des études non sélectionnées n'est pas publiée car le WWC ne diffuse pas cette information pour ce type de méta-analyse (communication personnelle)².

Le codage

Dans cette synthèse, seules les mathématiques étaient considérées. Chaque étude sélectionnée est associée à au moins une des 6 recommandations émises par les auteurs de la synthèse (voir encadré) et à au moins un des treize domaines mathématiques dont la liste complète a été définie *a priori* dans le protocole. La première recommandation (mettre en œuvre un enseignement explicite) concerne la façon d'enseigner, les 5 autres recommandations sont liées à des sujets, des types d'exercices et/ou des compétences mathématiques à développer chez les élèves. Chaque étude aboutie au calcul d'une ou plusieurs tailles d'effet, chaque taille d'effet étant toujours associée à un des 13 domaines mathématiques. Pour chacune des recommandations, les tailles d'effet globales pour chaque domaine mathématique sont ensuite calculées³. Aucune taille d'effet globale n'est calculée au niveau de la recommandation.

Les informations et les résultats concernant les études sélectionnées sont publiés dans six annexes (une par recommandation) :

- un premier tableau donne, pour chaque domaine mathématique, le nombre d'études associées, les tailles d'échantillon totales et les tailles d'effet globales calculées

¹ C'est l'année 2018 qui est mentionnée dans le protocole. Certaines études sont antérieures à 2004 car avaient été sélectionnées lors d'une première méta-analyse du WWC datée de 2009 et ont été conservées. Dans ce document, les études sont postérieures à 1983.

² Contrairement aux Rapports d'interventions qui sont des méta-analyses également publiées par le WWC.

³ C'est le modèle des effets fixes qui est utilisé ici, voir www.mathadoc.fr

- un second tableau détaille plus précisément chacune des études sélectionnées (design, description de l'intervention, taille des échantillons traitement et contrôle, tailles d'effet calculées par le WWC toujours associées à un domaine mathématique).

- **Recommandation 1** : Enseignement explicite - fournir un enseignement systématique durant les interventions pour développer la compréhension par l'élève des idées mathématiques.
- **Recommandation 2** : Langage mathématique - enseigner une langue mathématique claire et concise et soutenir l'utilisation par les élèves d'un langage leur permettant de communiquer efficacement leur compréhension des concepts mathématiques.
- **Recommandation 3** : Représentations - utiliser des représentations concrètes et semi-concrètes soigneusement sélectionnées pour faciliter la compréhension par les élèves de concepts et procédures mathématiques, élaborer un matériel adapté au niveau d'étude, et préparer les élèves à des mathématiques avancées.
- **Recommandation 4** : Droite numérique - utiliser une droite graduée (ou ligne des nombres) pour faciliter l'apprentissage par les élèves de concepts et procédures mathématiques.
- **Recommandation 5** : Problèmes verbaux - donner des instructions explicites sur les problèmes verbaux pour approfondir la compréhension mathématique des élèves et renforcer leur aptitude à appliquer des idées mathématiques.
- **Recommandation 6** : Activités chronométrées – inclure de façon régulière des activités chronométrées pour favoriser l'aisance en mathématique.

Résultats généraux

La principale conclusion du What Works Clearinghouse tient dans la liste des 6 recommandations qui sont associées aux études sélectionnées. Ces recommandations ne sont pas mentionnées dans le protocole et semblent donc avoir été définies *a posteriori*. Elles sont toutes évaluées comme ayant un niveau de preuve important¹ (le WWC définit trois niveaux de preuve (faible, modéré, important) en fonction de critères qui sont rappelés dans chacun de leur Guide des pratiques et qui sont détaillés dans le WWC Procedures Handbook 4.0).

¹ C'est la première fois qu'un Guide des Pratiques concernant l'enseignement des mathématiques affiche des recommandations qui obtiennent toutes ce niveau de preuve.

Les études sélectionnées

Le What Works Clearinghouse a sélectionné 44 études dont 41 ont été publiées entre 2004 et 2019¹. Les références des études sélectionnées sont écrites en gras dans la liste des références.

Pour clore cette rapide présentation des méthodes et des principaux résultats proposés par ces deux méta-analyses, on remarquera que les critères d'éligibilité ne comprennent pas de taille minimum requise pour les échantillons ni de durée de mise en œuvre de l'intervention minimale, comme c'est parfois le cas dans certaines synthèses de ce type (voir par exemple la méta-analyse de Robert SLAVIN (2009) qui fixait un seuil minimum de 15 élèves par groupes et une durée de l'intervention d'au moins 12 semaines).

Les études sélectionnées par Campbell et par le WWC

On va s'intéresser ici aux études sélectionnées par Campbell ou par le WWC en distinguant d'une part les études sélectionnées par les deux méta-analyses (les études communes), d'autre part les études sélectionnées par l'une des méta-analyses et non sélectionnées par l'autre. Dans ce dernier cas, il peut s'agir d'études non éligibles ou bien d'études éligibles mais dont la qualité méthodologique a été jugée insuffisante. Si certains critères d'éligibilité (comme la date de publication ou la langue) permettent de comprendre l'absence de certaines études dans une liste², d'autres motifs d'inéligibilité restent inconnus. Parmi les 58 études sélectionnées par Campbell (associées à au moins un domaine et à un test en mathématiques, voir ci-dessus), 16 ont été publiées avant 2004 ou dans une autre langue que l'anglais³, et ne pouvaient pas être sélectionnées par le What Works Clearinghouse. Parmi les 42 publications restantes, 18 ont également été sélectionnées par le WWC. Une étude (FUCHS, 2014a) n'a pas été sélectionnée par le WWC qui a par contre sélectionné une autre étude (POWELS, 2015) non retenue par Campbell car son échantillon était le même que celui de l'étude de FUCHS (2014a) qui lui a été préféré. Donc, ce sont finalement 19 études qui peuvent être considérées comme ayant été sélectionnées à la fois par Campbell et par le WWC. On a finalement 23 études publiées par Campbell qui n'ont pas été sélectionnées par le WWC sans qu'on puisse en déterminer la raison.

Parmi les 44 études publiées par le WWC, 6 publications sont postérieures à juin 2018 et ne pouvaient pas être sélectionnées par Campbell. Parmi les 38 études restantes, trois études ont été écartées par Campbell soit parce qu'elles comparaient des groupes traitement entre eux (JITENDRA, 2013a et JITENDRA, 2013b) soit car présentant un risque de biais important (SWANSON et LUSSIER, 2013); enfin 19 études (voir ci-dessus) ont également été

¹ Trois études ont été publiées avant 2004 car elles avaient été sélectionnées lors d'une première méta-analyse du WWC.

² Par exemple, le WWC n'a pas retenu les études analysant des interventions en classe entière et Campbell n'a pas retenu les études n'utilisant pas des tests standardisés.

³ Une des 58 études était écrite en suédois (HANSSON, 2014).

sélectionnées par Campbell. Il reste donc 16 études qui ne sont pas référencées par Campbell sans qu'on puisse en déterminer la raison (**schéma 3**).

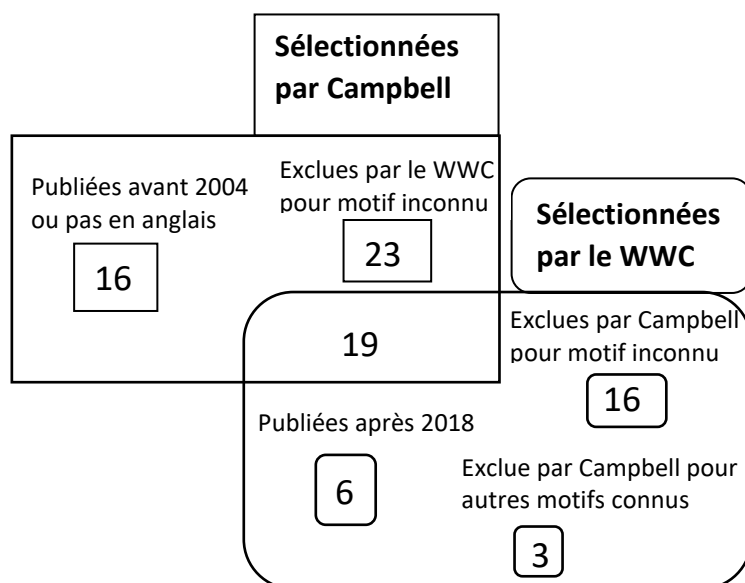


Schéma 3 : les études sélectionnées par Campbell et le WWC

Si on s'intéresse plus particulièrement aux études non sélectionnées par les deux méta-analyses pour des motifs qui n'ont pas pu être déterminés, on remarque qu'environ une étude sur deux est concernée (55% pour le WWC et 46% pour Campbell, **tableau 2**). Ainsi si 19 études ont été sélectionnées à la fois par le WWC et Campbell, 23 études n'ont pas été sélectionnées par le WWC pour un motif inconnu (mais sélectionnées par Campbell) ce qui représentent 55% de cet ensemble.

A	WWC		Campbell	
	nombre	%	nombre	%
Non sélectionnées par A pour motif inconnu	23	55%	16	46%
Sélectionnées par A	19	45%	19	54%
Total	42	100%	35	100%

Tableau 2 : études non sélectionnées pour motif inconnu

Les listes des études sélectionnées et non sélectionnées sont proposées dans l'annexe 2 (www.mathadoc.fr).

Le codage

Les textes intégraux des 19 études primaires communes aux deux méta-analyses ont été téléchargés à l'exception d'une étude (FUCHS, 2005) dont seul le résumé publié sur internet a été consulté. Le **tableau 3** indique les catégories associées à chacune d'entre elles. L'ensemble de ces 19 études sont caractérisées par un enseignement explicite en ce qui

concerne le WWC, à l'exception d'une étude (SMITH, 2013). Toutes ces études sont caractérisées par Campbell comme ayant mis en place un enseignement en petits groupes (maximum 6 élèves par groupe), à l'exception d'une étude (FIEN, 2016).

Auteurs	Campbell												WWC					
	Enseignement assisté par ordinateur	Formation professionnelle	Récompenses	Suivi des progrès	Enseignement en petits groupes	Algèbre	Fractions	Géométrie	Mathématiques générales	Connaissance des nombres	Operations	Résolution de problème	Enseignement explicite	Langage Mathématique	Représentations	Ligne des nombres	Problèmes verbaux	Activités chronométrées
Bryant (2011)				1	1				1	1	1		1		1			1
Clarke (2014)					1	1			1	1	1	1	1		1			
Dyson (2015)					1				1	1	1	1	1	1	1			1
Fien (2016)	1					1			1	1	1	1	1		1			1
Fuchs (2005)	1		1	1	1				1	1	1		1		1			1
Fuchs (2008a)*					1	1			1		1	1	1				1	1
Fuchs (2008b)					1						1	1					1	
Fuchs (2009)	1		1		1	1			1		1	1	1	1			1	1
Fuchs (2010)			1		1	1			1	1	1	1	1				1	1
Fuchs (2013a)					1		1						1	1	1	1		1
Fuchs (2013b)			1		1				1	1	1		1	1	1	1		1
Powels (2015)** WWC	-		-		-	-	-	-	-	-	-	-	1		1			1
Fuchs (2014a)** Campbell					1	1			1		1	1	-	-	-	-	-	-
Fuchs (2014b)					1		1		1		1		1	1	1	1		1
Fuchs (2016b)			1		1				1	1	1	1	1	1	1	1	1	1
Fuchs (2016c)			1		1		1		1	1		1	1	1	1	1	1	1
Gersten (2015)					1				1	1	1		1		1	1		1
Smith (2013)		1			1				1	1	1			1	1			
Swanson (2014a)					1	1		1	1			1	1				1	
Swanson (2014b)					1	1		1	1		1	1	1				1	

Tableau 3 : les 19 études primaires communes à Campbell et au WWC

* : non téléchargée

** : ces deux publications sont considérées comme issues de la même étude

En **gras** les caractéristiques associées à des méthodes d'enseignement, en *italique* les caractéristiques associées à des domaines ou à des compétences mathématiques. Les références sont données sous le format utilisé par Campbell.

On va s'intéresser aux 17 études associées à un enseignement explicite par le WWC et à un enseignement en petits groupes par Campbell et porter notre attention sur le ou les groupes traitement ainsi que le groupe contrôle analysés dans chaque étude afin de comprendre

quelles comparaisons ont été menées pour préciser les interventions évaluées (voir annexe 1 p.27 et annexe 2 sur www.mathadoc.fr).

Pour 15 études, le groupe contrôle était constitué d'élèves à risque ou en difficulté en mathématiques recevant un enseignement en classe entière habituel et le ou les groupes traitements étaient constitués d'élèves à risque ou en difficulté en mathématiques bénéficiant de séances de soutien (soit individuel, soit en petits groupes). Ces séances sont généralement réparties en 3 ou 4 périodes hebdomadaires de 30 minutes pendant au moins 12 semaines, en dehors des cours en classe entière, et sont la plupart du temps dénommées *tutoring* par les auteurs des études primaires. L'enseignement pour les groupes traitement est décrit avec précision (plusieurs méthodes ou programmes différents ont été parfois mis en œuvre dans plusieurs groupes pour une même étude) et cet enseignement peut être caractérisé comme explicite. A l'inverse, la plupart du temps l'enseignement du groupe contrôle en classe entière n'est que peu décrit ; dans un cas il est mentionné que cet enseignement est explicite (FUCHS, 2010), dans un autre il est mentionné que l'enseignement n'est pas explicite (BRYANT, 2011)¹.

En fait, ces études ont analysé la mise en place d'un enseignement basé sur les concepts de la Réponse à l'intervention (*Response to Intervention, Rti*) souvent mis en œuvre de nos jours aux Etats-Unis et dont une rapide description est donnée dans l'encadré ci-dessous.

Réponse à l'intervention

La caractéristique majeure de cette approche consiste à proposer trois niveaux d'enseignement :

- un premier niveau (*Tier 1*) avec des cours en classe entière où une attention particulière est portée aux élèves en difficulté repérés lors d'évaluations périodiques et où l'enseignement dispensé est fondé sur des preuves scientifiques
- un deuxième niveau (*Tier 2*) avec des séances de soutien supplémentaire dispensées en petits groupes pour les élèves qui ne progressent pas suffisamment au premier niveau
- un troisième niveau (*Tier 3*) avec des séances de soutien spécialisé (en petits groupes ou individualisé) parfois pendant les heures de classe entière pour les élèves qui ont des troubles d'apprentissage et/ou qui ne progressent pas au deuxième niveau.

Cette organisation de l'enseignement a comme objectif d'améliorer les compétences et connaissances académiques des élèves.

On parle aussi de *multi-tier system of supports* (MTSS)².

¹ La prédominance des interventions visant à évaluer l'impact d'une forme de *tutoring* et classées dans la catégorie « enseignement en petits groupes » est mentionnée dans la synthèse de Campbell (DIETRICHSON, 2021, p.52).

² Une formation des professionnels est alors souvent incluse et des compétences comportementales sont également ciblées.

Pour terminer la rapide présentation de ces 15 études, on pourra remarquer qu'elles ne comparent *pas* de façon formelle :

- un enseignement en classe entière où les élèves travaillent en petits groupes avec un enseignement en classe entière où les élèves ne travaillent pas en petits groupes
- un enseignement explicite avec un enseignement non explicite.

Pour les 2 études restantes (SWANSON, 2014a et SWANSON, 2014b), le groupe contrôle était constitué d'élèves à risque ou en difficulté en mathématiques recevant un soutien en petits groupes en dehors des heures de classe entière ; les élèves à risque ou en difficulté en mathématiques des groupes traitement recevaient également un soutien en petits groupes en dehors des heures de classe entière avec des pratiques pédagogiques différentes où l'enseignement explicite est clairement mentionné.

Les tailles d'effet des études primaires

Les deux méta-analyses ont calculé et publié des tailles d'effet pour chacune des études primaires sélectionnées. Ces deux synthèses ont choisi le g de Hedges qui, au moins dans les cas simples où ce sont les élèves (et non des classes ou des établissements) qui sont attribués aléatoirement dans le groupe contrôle ou le groupe traitement, devraient donner les mêmes résultats. Pour les 17 études identifiées précédemment, Campbell a calculé 68 tailles d'effet et le WWC en a calculé 59. Ces tailles d'effet ont été représentées graphiquement (**schéma 4**) (chaque étude a été identifiée par un numéro, voir l'annexe 3 pour plus de précisions).

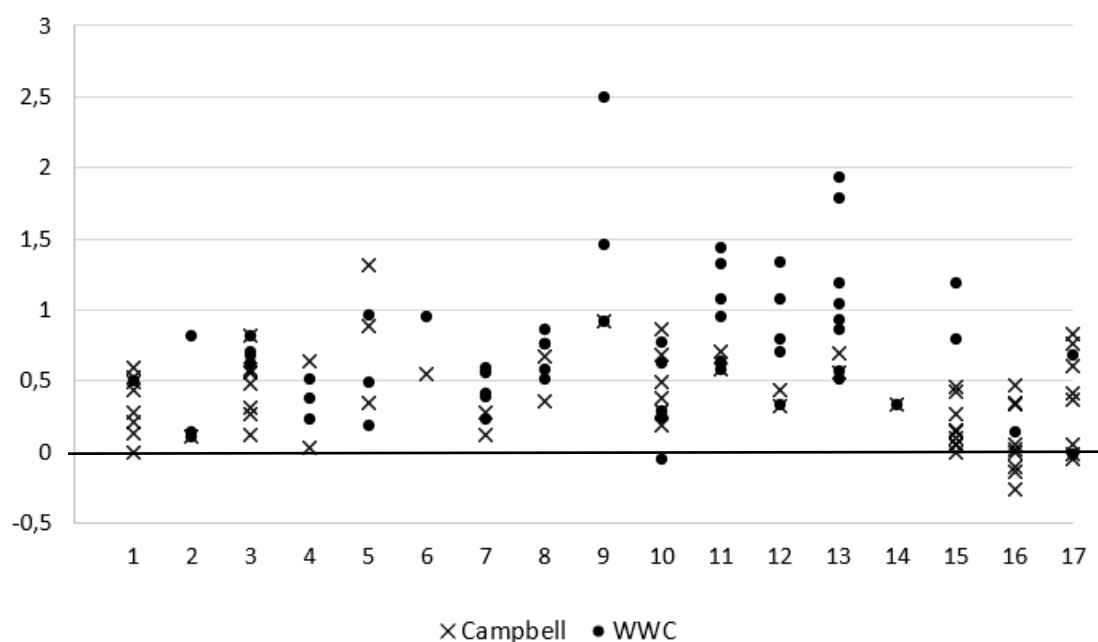


Schéma 4 : tailles d'effet des études primaires sélectionnées par Campbell et le WWC.

Pour chaque taille d'effet calculée, le WWC mentionne si elle est (ou non) statistiquement significative au niveau de confiance de 0,95¹. Dans la méta-analyse de Campbell, des erreurs standards sont publiées pour chacune des tailles d'effet ce qui a permis de calculer une valeur-p (toujours pour un niveau de confiance de 0,95) en utilisant un test-z. Les résultats sont présentés en annexe 3 et sont résumés dans le **tableau 4**. Ainsi, 46% des tailles d'effet calculées par Campbell sont positives et statistiquement significatives, et cette proportion est de 85 % pour le WWC.

		Campbell			WWC		
		Nombre	Nombre total	%	Nombre	Nombre total	%
ES>=0	p> 0,05	30	61	44%	7	57	12%
	p< 0,05	31		46%	50		85%
ES<0	p> 0,05	7	7	10%	2	2	3%
	p< 0,05	0		0%	0		0%

Tableau 4 : caractéristiques des tailles d'effet (ES²) calculées par Campbell et le WWC

La proportion des tailles d'effet statistiquement significatives est plus importante pour le WWC. On remarquera que le WWC calcule souvent des tailles d'effet en regroupant plusieurs échantillons traitement différents, ce qui a comme effet de diminuer les variances des tailles d'effet et donc de diminuer les valeurs-p (ce qui peut conduire à une élévation de la significativité statistique des résultats). C'est aussi ce qui explique les différences entre les tailles d'effet calculées entre les deux méta-analyses pour une même étude.

La conclusion principale que l'on peut sans doute tirer de ces 17 études primaires est que l'impact sur les résultats en mathématiques des élèves en difficulté (ou à risque de l'être) de séances de soutien fréquentes en petits groupes en dehors des heures d'enseignement en classe entière (*tutoring*) et mettant en œuvre un enseignement que l'on peut qualifier d'explicite, est vraisemblablement positif.

L'apprentissage soutenu par les pairs

Dans tout ce qui suit, les résultats sont à consulter dans l'annexe 4.

Au début de ce texte, il avait été souligné que l'impact de l'apprentissage soutenu par les pairs (*peer-assisted instruction*) est présenté par la synthèse de Campbell comme un résultat remarquable³ avec une taille d'effet globale égale à 0,44 (CI = [0,28 ; 0,61]). Aucune des 19 études communes aux deux méta-analyses n'est associée à cette caractéristique. Pour comprendre cette absence, les études sélectionnées par Campbell ont été filtrées en ne conservant que les études associées à l'apprentissage soutenu par les pairs et qui ont utilisé des tests mathématiques. Elles sont au nombre de 14. Treize d'entre elles ont été publiées avant 2004, ce qui explique bien évidemment qu'elles n'aient pas pu être sélectionnées par le WWC et qu'elles ne figurent pas dans la liste des études communes au WWC et à

¹ Ce qui revient à dire que la valeur-p est (ou non) inférieure à 0,05.

² ES = tailles d'effet (*Effect Size*)

³ Rappelons que ce résultat concerne globalement les apprentissages en mathématiques et en lecture.

Campbell. Il reste une étude (NUSSBAUM, 2010) dont on ne peut pas déterminer la cause de non-sélection par le WWC.

Dans ce qui suit, seuls les résultats de la méta-analyse de Campbell ont été considérés. On s'est attaché à comparer les études et les tailles d'effet calculées à partir de leurs résultats à des tests en mathématiques, quand elles analysaient des interventions ciblées soit sur l'apprentissage soutenu par les pairs, soit sur l'enseignement en petits groupes. Une seule étude (SLAVIN, 1984a) est associée à un enseignement en petits groupes et à l'apprentissage soutenu par les pairs.

Le nombre d'études associées à chacune de ces 2 catégories ainsi que la moyenne de l'année de leur publication et la somme des tailles d'échantillons totales, sont présentées dans le **tableau 5** ci-dessous. La taille d'échantillon totale d'une étude est la plus grande somme de l'effectif du groupe traitement et de l'effectif du groupe contrôle que l'on peut calculer pour cette étude (certaines études analysent plusieurs échantillons différents). Les résultats sont représentés graphiquement pour le nuage de points ci-dessous (**schéma 5**) où chaque point représente une étude avec en abscisse son année de publication et en ordonnée la taille d'échantillon totale définie auparavant. On remarque que les résultats des études portant sur l'enseignement en petits groupes sont plus récents, ont été menés sur des d'échantillon de taille plus importante et ont concernés beaucoup plus d'élèves que les résultats des études portant sur l'apprentissage soutenu par les pairs. Une des études associées à l'enseignement en petits groupes (BERNSTEIN, 2009) n'a pas été représentée car la taille de son échantillon importante¹ écrasait la représentation graphique.

	Soutien par des pairs	Enseignement en petits groupes
Nombre d'études	14	39
Moyenne année publication	1993,9	2010,8
Somme des tailles d'échantillon totales	912,5	10 198
		7 838*

Tableau 5 : caractéristiques des études associées à l'enseignement en petits groupes et à l'apprentissage soutenu par les pairs

* : valeurs calculées en excluant l'étude de Bernstein (2009) dont l'effectif peut être considéré comme anormalement élevé (et est non représenté dans le **schéma 5**).

¹ Cette taille d'échantillon est égale à 2 360.

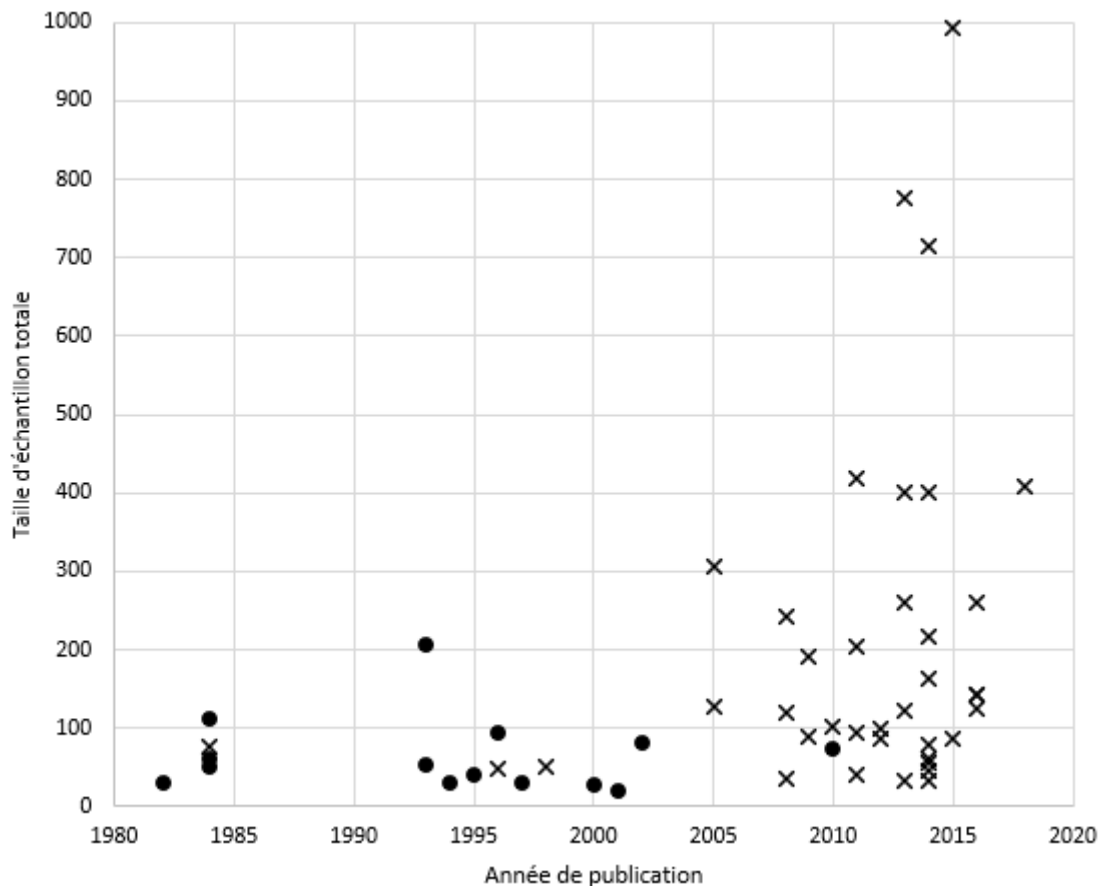


Schéma 5 : études associées à l’enseignement en petits groupes (x) et associées à l’apprentissage soutenu par les pairs (•).

Les tailles d’effet (ES) des études associées à ces deux catégories ont également été caractérisées. Pour les 52 études identifiées précédemment¹, Campbell a calculé 32 tailles d’effet associées à l’apprentissage soutenu par les pairs et 147 tailles d’effet associées à l’enseignement en petits groupes. Les erreurs standards de ces tailles d’effet étant publiées, il a été possible de calculer une valeur-p au niveau de confiance de 0,95 en utilisant un test z. Le nombre moyen d’observations pour chacune de deux catégories est également présenté. Ce nombre est calculé à partir de toutes les observations de chaque étude ; dans une même étude, plusieurs tailles d’effet ont pu être calculées avec soit des groupes identiques (quand plusieurs tests sont effectués), soit des groupes différents (quand plusieurs traitements sont étudiés). Les notions d’*observation* (mesure d’un score) et de *sujet* (ici un élève) ont été distinguées (**tableau 6**).

Ainsi 16% des tailles d’effet calculées associées à l’apprentissage soutenu par les pairs sont positives et statistiquement significatives, et cette proportion est égale à 42% pour l’enseignement en petits groupes. On remarque aussi que le nombre d’observations

¹ Une étude est associée à l’apprentissage soutenu par les pairs et l’enseignement en petits groupes (SLAVIN, 1984a).

analysées dans les études sur l'enseignement en petits groupes est en moyenne¹ supérieure au nombre d'observations analysées dans les études sur l'apprentissage soutenu par les pairs (ce qui explique en partie la différence entre ces deux types d'études de la proportion de résultats statistiquement significatifs au niveau de confiance 0,95). On notera que la valeur-p calculée à partir d'un test-z sur les 32 tailles d'effet associées à l'apprentissage soutenu par les pairs est inférieure à 0,05 pour 5 tailles d'effet (toutes positives).

		Soutien par les pairs			Petits groupes					
		Nombre	Total	%	Nombre	Total	%	Nombre *	Total*	%*
ES>=0	p> 0,05	20	25	63%	69	131	47%	69	129	48%
	p< 0,05	5		16%	62		42%	60		41%
ES<0	p> 0,05	7	7	22%	16	16	11%	16	16	11%
	p< 0,05	0		0%	0		0%	0		0%
Nombre d'observations moyen		50,9			205,4			207,4		

Tableau 6 : caractéristiques des tailles d'effet (ES) associées à l'enseignement en petits groupes et à l'apprentissage soutenu par les pairs

* : Valeurs calculées sans les deux tailles d'effet supérieures à 3 considérées comme aberrantes (étude n°25, KIDD, 2014) et non représentée dans le **schéma 7**.

On a représenté graphiquement l'ensemble de ces tailles d'effet pour chacune des deux catégories (**schéma 6** pour l'apprentissage soutenu par les pairs, **schéma 7** pour l'enseignement en petits groupes).

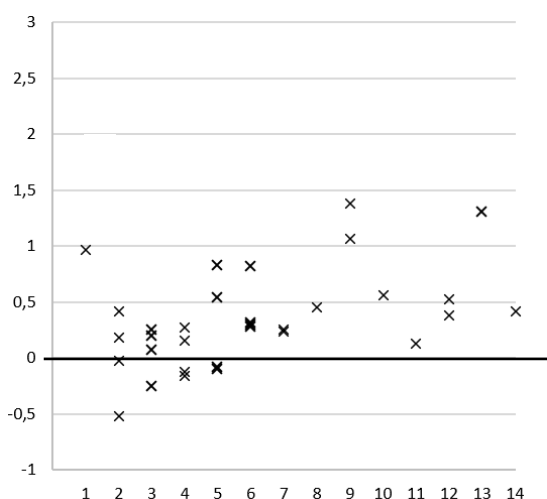


Schéma 6 : 32 tailles d'effets des études associées à l'apprentissage soutenu par les pairs

¹ On calcule la somme des observations divisée par leur effectif (c'est -à-dire le nombre de tailles d'effet calculées).

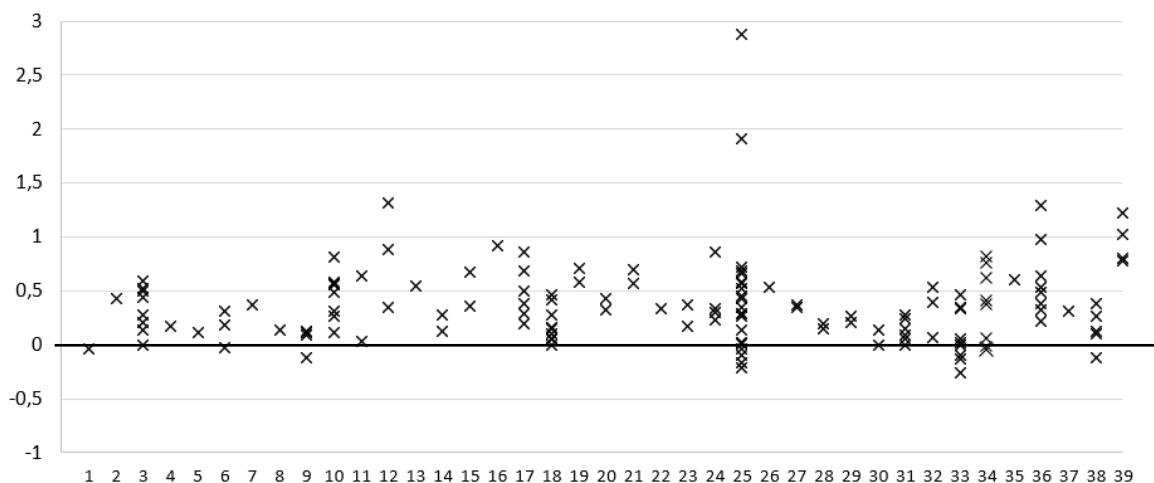


Schéma 7 : 147 tailles d'effets des études associées à l'enseignement en petits groupes. Deux tailles d'effet supérieures à 3 n'ont pas été reportées (étude n°25, KIDD, 2014).

Pour conclure

Deux méta-analyses contemporaines l'une de l'autre et partageant une même question ont finalement répondu de façons différentes. Comme nous l'avons montré, dès la première étape de ces méta-analyses qui consiste à sélectionner des études, les résultats divergent : 19 études sont à la fois sélectionnées par le WWC et Campbell, mais presque autant d'études sont laissées de côté par chacun des deux organismes sans que l'on puisse toujours savoir pourquoi. Ensuite, si on considère uniquement les 19 études communes, les résultats et les conclusions proposées divergent là encore : leurs codages sont différents et mettent en valeur des caractéristiques différentes ; et enfin les tailles d'effet calculées sont également différentes y compris pour une même étude.

Des différences qui peuvent s'expliquer

C'est dans le détail des protocoles que se logent les divergences, et par exemple un critère d'éligibilité aussi anodin qu'une date en deçà de laquelle les études ne sont pas sélectionnées peut changer la donne (comme nous l'avons vu en ce qui concerne l'apprentissage soutenu par les pairs par exemple).

Les différentes façons d'envisager la question de recherche (et les réponses attendues ou tout du moins compatibles avec une certaine vision de l'enseignement et des apprentissages) orientent l'attention du chercheur sur des caractéristiques qui lui semblent pertinentes. Les études primaires seront alors codées à l'aide de ce prisme de lecture qui peut ne pas être le même d'un chercheur à l'autre. Et il en va de même quand on considère les objectifs. Le Guide des pratiques du WWC, consacré uniquement aux mathématiques, annonce dès son introduction que le soutien supplémentaire (ou encore *tutoring*) mis en place pour les élèves en difficulté ne relève plus de la question de recherche car son

efficacité n'est plus à prouver¹. La volonté de proposer aux enseignants des solutions concrètes sur la manière optimale d'organiser ces séances et de mettre à leur disposition des outils (comme par exemple des exercices-type) à utiliser avec des élèves en difficulté en mathématiques, a clairement orienté la méta-analyse sur un terrain concret. La méta-analyse de Campbell s'inscrit dans un cadre scientifique rigoureux et propose de très nombreux résultats statistiques qui n'ont pas été évoqués ici ; les auteurs n'avaient pas comme objectif de proposer des recommandations détaillées aux enseignants mais plutôt d'identifier les caractéristiques des interventions qui semblent être efficace. On s'étonnera tout de même de l'absence de l'enseignement explicite de leur liste des caractéristiques pédagogiques.

Les calculs des tailles d'effet sont également soumis à des variations quand on rentre dans le détail des formules : certains cas complexes (comme des assignations aléatoires au niveau des classes et non des élèves, ou des mesures multiples) sont traités de manière légèrement différente². Enfin quand plusieurs groupes traitement sont analysés dans une étude, ou que plusieurs tests sont mis en œuvre, il est alors possible de calculer un nombre important de tailles d'effet. C'est alors la préférence pour tel type de test ou pour tel regroupement possible de résultats³, qui peut expliquer que des résultats différents sont proposés pour une même étude. La méthode statistique employée par Campbell utilise notamment le modèle des effets aléatoires conjointement à l'estimation d'une variance robuste (*robust-variance estimation*) qui permet de tenir compte de la dépendance entre les tailles d'effet d'une même étude.

Les questions à se poser

Le choix des caractéristiques que l'on associe à chaque étude primaire est fondamental, notamment quand ces caractéristiques concernent des méthodes pédagogiques, car c'est vers elles que se tournent tous les regards. Pour y parvenir, une description fine des traitements mis en œuvre dans le ou les groupes traitement mais aussi des enseignements mis en œuvre dans les groupes contrôle souvent simplement qualifiés d'« habituels » (*business as usual*) est indispensable. En effet, c'est elle qui permettra de comprendre ce qui différencie les expériences vécues par les élèves selon qu'ils appartiennent à un groupe ou à un autre dans le but d'identifier, de définir et de caractériser les *différences* entre ces groupes (et non pas simplement l'intervention en tant que telle), et regrouper les études primaires de façon cohérente et objective en fonction des résultats de cette analyse.

¹ Parmi les 19 études communes, la seule qui ne soit pas associée par Campbell à l'enseignement en petits groupes (FIEN, 2016) a comparé deux groupes d'élèves qui suivaient des séances de soutien (tier 2) et l'intervention consistait en l'utilisation par les élèves d'un logiciel déployé sur ordinateur basé sur les concepts de l'enseignement explicite.

² Par exemple les coefficients intra-classe utilisés sont parfois différents.

³ Campbell n'a pas retenu les résultats issus de tests « maison » et le WWC a souvent procédé à des regroupements.

Comme nous l'avons déjà évoqué, d'autres méta-analyse excluent les études dont les échantillons n'atteignent pas une taille suffisante. En effet, il a été montré que les tailles d'effet des petites études étaient souvent surestimées, et une attitude conservatrice peut donc amener à ne pas considérer leurs résultats¹. On remarquera également que les tailles d'effet de ces études de petite taille sont moins souvent statistiquement significatives par manque de puissance.

Enfin les auteurs de ces méta-analyses soulignent dans leurs articles que plusieurs facteurs peuvent expliquer les résultats d'une seule et même intervention, facteurs qui peuvent parfois interagir entre eux ou interagir avec l'intervention². Les interventions étudiées sont également souvent associées à plusieurs caractéristiques pédagogiques et dans ce cas il n'est pas possible d'identifier l'une ou l'autre (ou une de leurs combinaisons) comme responsable (ou non) de l'effet mesuré. C'est pour cette raison que les auteurs de la méta-analyses de Campbell n'évoquent pas de lien de cause à effet entre une caractéristique et un résultat mais bien une association³.

¹ La qualité des études de petite taille est souvent considérée comme plus basse que la qualité des études de grande envergure. Ce point a été commenté dans la méta-analyse de Campbell (p. 45)

² Une analyse de l'effet de certains modérateurs a été menée par Campbell en utilisant un modèle de régression linéaire multiple et une analyse de sous-groupe a également été conduite.

³ C'est tout particulièrement vrai pour les analyses exploratoires qui sont menées au vu des premiers résultats obtenus et qui ne peuvent de ce fait prétendre à une validité interne telle qu'elle est attendue dans les analyses qui ont comme objectif de confirmer (ou d'infirmer) une hypothèse. Un autre problème se pose également quand de nombreuses analyses statistiques sont menées, car cette multiplicité conduit à une inflation du risque d'erreur de type I dont il faudrait tenir compte (ce problème est évoqué dans la méta-analyse de Campbell, p.35).

Glossaire

Caractéristique pédagogique	Élément descriptif caractéristique d'une méthode d'enseignement.
Codage	Identification et recension des éléments descriptifs d'une étude de son (ses) intervention(s), de sa (ses) tailles d'effet : noms d'auteurs et année de publication pour une étude, caractéristiques pédagogiques ou domaines pour une intervention, taille d'échantillon ou nature du test pour une taille d'effet.
Domaine	Notion(s) ou concept(s) qui sont visés par l'enseignement.
Éligible	Une étude est éligible quand elle est conforme aux critères d'éligibilité (population cible, type d'intervention prédéfinie, design de l'étude, année de publication, langue, secteurs géographiques par exemple).
Étude primaire	Étude expérimentale prospective dont proviennent les données utilisées par les méta-analyses.
Intervention	Cela peut être un programme scolaire ou un ensemble de pratiques qualifié de méthode pédagogique ou un produit commercial (comme une application informatique ou un manuel), dont l'objectif est d'améliorer les compétences des élèves. On parle aussi de traitement.
Méta-analyse	Synthèse quantitative d'études primaires sélectionnées utilisant des procédures statistiques.
Sélection	Opération visant à inclure (ou rejeter) des études dans une méta-analyses selon des critères prédéfinis, dans l'objectif d'assurer la validité interne des conclusions en confirmant le lien de cause à effet du traitement étudié.
Taille d'effet	Différence standardisée des moyennes : écart entre les moyennes divisé par l'écart-type des scores le g de Hedges ici)
Validité interne	Qualité d'une étude associée à la fiabilité des conclusions internes (lien de cause à effet)

Abréviations

ES	<i>Effect size</i> traduit par taille d'effet
WWC	What Works Clearinghouse

Annexes

Annexe 1 : Description des 19 études communes à Campbell et au WWC (p.27)

Annexe 2 : Les publications de Campbell et du WWC (www.mathadoc.fr)

Annexe 3 : Les résultats de Campbell et du WWC (www.mathadoc.fr)

Annexe 4 : Les résultats de Campbell (www.mathadoc.fr)

Références

Certaines références sont à consulter dans les méta-analyses d'origine.

*Les études précédées du symbole * sont communes à Campbell et au WWC.*

*Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early numeracy intervention program for first-grade students with mathematics difficulties. *Exceptional Children*, 78(1), 7–23. <https://eric.ed.gov/?id=EJ939950>

Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45 (5), 283-292.

*Clarke, B., Doabler, C. T., Strand Cary, M., Kosty, D., Baker, S., Fien, H., & Smolkowski, K. (2014). Preliminary evaluation of a tier 2 mathematics intervention for first-grade students: Using a theory of change to guide formative evaluation activities. *School Psychology Review*, 43(2), 160–178. <https://eric.ed.gov/?id=EJ1142175>

Cooper, A., Hedges, L. & Valentine, J. (2019). *The Handbook of Research Synthesis and Meta-Analysis*. NEW YORK: Russell Sage Foundation.

Dietrichson, J., Filges, T., Seerup, J. K., Klokke, R. H., Viinholt, B. C. A., Bøg, M., & Eiberg, M. *Targeted school - based interventions for improving reading and mathematics for students with or at risk of academic difficulties in Grades K - 6: A systematic review*. *Campbell Systematic Reviews*. 2021;e1152. <https://doi.org/10.1002/cl2.1152>

*Dyson, N., Jordan, N. C., Beliakoff, A., & Hassinger-Das, B. (2015). A kindergarten number-sense intervention with contrasting practice conditions for low-achieving children. *Journal for Research in Mathematics Education*, 46(3), 331–370. <https://eric.ed.gov/?id=EJ1088271>

*Fien, H., Doabler, C. T., Nelson, N. J., Kosty, D. B., Clarke, B., & Baker, S. K. (2016). An examination of the promise of the NumberShire Level 1 gaming intervention for improving student mathematics outcomes. *Journal of Research on Educational Effectiveness*, 9 (4), 635–661. <https://eric.ed.gov/?id=ED576652>

*Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, 97(3), 493–513. <https://eric.ed.gov/?id=EJ734286>

*Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008 a). Effects of preventative *tutoring* on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, 74(2), 155–173. <https://eric.ed.gov/?id=EJ817525>

*Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., & Hamlett, C. L. (2008 b). Effects of small-group *tutoring* with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*, 100(3), 491–509. <https://eric.ed.gov/?id=EJ807859>

*Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., Hamlett, C. L., & Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology*, 101(3), 561–576. <https://eric.ed.gov/?id=EJ861181>

*Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences*, 20 (2), 89–100. <https://eric.ed.gov/?id=EJ872585>

*Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013 a). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, 105(3), 683–700. <https://eric.ed.gov/?id=EJ1054396>

*Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., DeSelms, J., Seethaler, P. M., Wilson, J., Craddock, C. F., Bryant, J. D., Luther, K., & Changas, P. (2013 b). Effects of first-grade number knowledge *tutoring* with contrasting forms of practice. *Journal of Educational Psychology*, 105(1), 58–77. <https://eric.ed.gov/?id=EJ1006608>

*Fuchs, L. S., Powell, S. R., Cirino, P. T., Schumacher, R. F., Marrin, S., Hamlett, C. L., Fuchs, D., Compton, D. L., & Changas, P. C. (2014 a). Does calculation or word - problem instruction provide a stronger route to realgebraic knowledge? *Journal of Educational Psychology*, 106(4), 990 – 1006.

*Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., Hamlett, C. L., Jordan, N. C., Gersten, R., Siegler, R. S., & Changas, P. (2014 b). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, 106(2), 499–514. <https://eric.ed.gov/?id=EJ1054465>

*Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Malone, A. S., Wang, A., Hamlett, C. L., Jordan, N. C., Siegler, R. S., & Changas, P. (2016 b). Effects of intervention to improve at-risk fourth graders' understanding, calculations, and word problems with fractions. *The Elementary School Journal*, 116(4), 625–651. <https://eric.ed.gov/?id=EJ1103953>

*Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., Hamlett, C. L., Jordan, N. C., Siegler, R. S., Gersten, R., & Changas, P. (2016 c). Supported self-explaining during fraction intervention. *Journal of Educational Psychology*, 108(4), 493–508. <https://eric.ed.gov/?id=EJ1099301>

Fuchs, L.S., Newman-Gonchar, R., Schumacher, R., Dougherty, B., Bucka, N., Karp, K.S., Woodward, J., Clarke, B., Jordan, N. C., Gersten, R., Jayanthi, M., Keating, B., and Morgan, S. (2021). *Assisting Students Struggling with Mathematics: Intervention in the Elementary Grades* (WWC 2021006). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://whatworks.ed.gov/>.

*Gersten, R., Rolfhus, E., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2015). Intervention for first graders with limited number knowledge: Large-scale replication of a randomized controlled trial. *American Educational Research Journal*, 52(3), 516–546. <https://eric.ed.gov/?id=EJ1063565>

*Powell, S. R., Fuchs, L. S., Cirino, P. T., Fuchs, D., Compton, D. L., & Changas, P. C. (2015). Effects of a multitier support system on calculation, word problem, and prealgebraic performance among at-risk learners. *Exceptional Children*, 81(4), 443–470. <https://eric.ed.gov/?id=EJ1065063>

Slavin, R.E., Lake, C., et Groff, C. (2009). Effective programs in middle and high school mathematics : A best-evidence synthesis. *Review of Educational Research*, 79 (2), 839-911.

*Smith, T. M., Cobb, P., Farran, D. C., Cordray, D. S., & Munter, C. (2013). Evaluating math recovery: Assessing the causal impact of a diagnostic *tutoring* program on student achievement. *American Educational Research Journal*, 50(2), 397–428. <https://eric.ed.gov/?id=EJ1005742>

*Swanson, H. L., Moran, A., Lussier, C., & Fung, W. (2014 a). The effect of explicit and direct generative strategy training and working memory on word problem-solving accuracy in children at risk for math difficulties. *Learning Disability Quarterly*, 37(2), 111–122. <https://eric.ed.gov/?id=EJ1022817>

*Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2014 b). The effects of mathematics strategy instruction for children with serious problem-solving difficulties. *Exceptional Children*, 80(2), 149–168. <https://eric.ed.gov/?id=EJ1048482>

Annexe 1 : Description des 19 études communes à Campbell et au WWC

	Contrôle	Traitement 1	Traitement 2	Remarques
Bryant et al. (2011)	Élève en classe entière, pas d'enseignement explicite	<i>Tutoring</i> , groupes de 3-5 élèves, 4 séances de 25 minutes hebdomadaires pendant 19 semaines, enseignement explicite, représentations visuelles		
Clarke (2014)	Élève en classe entière	<i>Fusion instruction</i> , nombres entiers, Tier 2, groupes de 5 élèves, 60 leçons de 30 minutes sur 20 semaines. Enseignement explicite.		
Dyson (2015)	Élève en classe entière	Groupes de 4 élèves, 24 sessions de 30 minutes (variante 1)	Groupes de 4 élèves, 24 sessions de 30 minutes (variante 2)	
Fien (2016)	Élève en classe entière + interventions supplémentaires (Tier 2)	Tier 2, intervention NS1, 4 séances de 15 minutes hebdomadaires pendant 8 semaines, enseignement explicite ; élève en groupes de 5 à 25.		Rmq : seule étude non caractérisée par un enseignement en petits groupes par Campbell, mais CAI.
Fuchs (2005) (non téléchargée)	Pas de <i>tutoring</i>	<i>Tutoring</i> 3 séances hebdomadaires pendant 16 semaines		
Fuchs (2008a)	Élève en classe entière, Tier 1, enseignement explicite mais pas sur les résolutions de problèmes	Tier 2, <i>tutoring</i> (1 : 1), 3 séances de 20-30 minutes hebdomadaires pendant 12 semaines sur la résolution de problèmes avec enseignement explicite		
Fuchs (2008b)	2 types d'enseignement en classe entières (classes conventionnelles ou classes SBI), enseignement explicite	<i>Tutoring</i> , sessions de 20-30 min, 940 minutes en tout		Rmq : 2 groupes traitements pourraient être considérés
Fuchs (2009)	Élève en classe entière, enseignement explicite	<i>Tutoring</i> , Maths Flash, SBI, 3 sessions de 20-30 minutes hebdomadaires pendant 16 semaines	<i>Tutoring</i> , Pirate Flash, SBI, 3 sessions de 20-30 minutes hebdomadaires pendant 16 semaines	
Fuchs (2010)	Élève en classe entière, enseignement explicite	<i>Tutoring</i> (1 : 1), 3 séances de 20-30 minutes hebdomadaires pendant 16 semaines, Pirate maths, résolution de problèmes, 1 seule leçon comptage stratégique	<i>Tutoring</i> (1 : 1), 3 séances de 20-30 minutes hebdomadaires pendant 16 semaines, Pirate maths, résolution de problèmes, plusieurs leçons comptage stratégique	
Fuchs (2013a)	Élève en classe entière. Interpréter les fractions comme des partages, procédures	Groupes de 3 élèves, interpréter les fractions comme des mesures, 3 fois 30 minutes hebdomadaires pendant 12 semaines.		

Fuchs (2013b)	Élève en classe entière	<i>Tutoring</i> (1:1) et <i>Speeding practice</i> (les 5 dernières minutes), 3 séances de 30 minutes hebdomadaires pendant 16 semaines	<i>Tutoring</i> (1:1) sans <i>Speeding practice</i> , 3 séances de 30 minutes hebdomadaires pendant 16 semaines	
Fuchs (2014a) Campbell Powels (2015) WWC	Élève en classe entière, 2 leçons de 40-45 min par semaines pendant 17 semaines	<i>Tutoring</i> calcul, groupes de 2-3 élèves, 3 séances de 25-30 minutes hebdomadaires pendant 13 semaines, enseignement explicite	<i>Tutoring</i> problèmes, groupes de 2-3 élèves, 3 séances de 25-30 minutes hebdomadaires pendant 13 semaines, enseignement explicite	
Fuchs (2014b)	Élève en classe entière. Interpréter les fractions comme des partages	Groupes de 3 élèves, <i>Tutoring</i> , 3 séances de 30 minutes hebdomadaires pendant 12 semaines dont 5 min <i>fluency</i>	Groupes de 3 élèves, <i>Tutoring</i> , 3 séances de 30 minutes hebdomadaires pendant 12 semaines dont 5 min élève explique son raisonnement	
Fuchs (2016b)	Élève en classe entière	Groupes de 2 élèves, 3 séances de 35 min hebdomadaires pendant 12 semaines sur les fractions (variante 1)	Groupes de 2 élèves, 3 séances de 35 min hebdomadaires pendant 12 semaines sur les fractions (variante 2)	
Fuchs (2016c)	Élève en classe entière. Séances supplémentaires en petits groupes de 27 minutes par semaine	Groupes de 2 élèves, 36 séances de 35 minutes (105 minutes par semaine pendant le cours en classe entière), comparaison de fractions, enseignement explicite	Groupes de 2 élèves, 36 séances de 35 minutes (105 minutes par semaine pendant le cours en classe entière), problèmes avec des fractions, enseignement explicite	
Gersten (2015)	Élève en classe entière	Groupes de 2-3 élèves, 30h en plus de la classe. Formation des intervenants		
Smith (2013)	Élève en classe entière	<i>Tutoring</i> (1 : 1), 4 ou 5 séances de 30 min hebdomadaires pendant 12 semaines		Rmq : seule étude non caractérisée par un enseignement explicite par le WWC.
Swanson (2014a)	Soutien classique	Soutien particulier 1 <i>Restate</i> , enseignement explicite	Soutien particulier 2 <i>Relevant</i> , enseignement explicite	+ deux autres groupes traitement, enseignement explicite
Swanson (2014b)	Soutien classique	Soutien particulier 1, enseignement explicite	Soutien particulier 2, enseignement explicite	+ deux autres groupes traitement, enseignement explicite

NS1 : NumberShire Level 1 Gaming Intervention ; SBI : Schema broadening instruction ; CAI : Computer Assisted Instruction