

EducMap : une carte scientométrique de la recherche en éducation vue par la base de données internationales *Scopus*

Kristine Lund, UMR 5191, CNRS, ENS-Lyon, Université de Lyon 2, France
Heisawn Jeong, Département de psychologie, Hallym University, Corée du Sud
Sebastian Grauwin, Université de Lyon, France
Pablo Jensen UMR 5672, CNRS, ENS-Lyon, France

Abstract: La recherche en éducation couvre un territoire de divers sujets ce qui présente un défi d'intégration. Nous avons effectué une analyse bibliographique de la base de données internationale *Scopus* afin de décrire la recherche autour du fait éducatif. Nous avons construit une carte globale — essentiellement de la recherche anglophone — consistant en dix-huit clusters distincts connectés par des ensembles de ressources bibliographiques. Les clusters diffèrent dans leurs références partagées et dans la façon dont ils se focalisent sur un aspect de la recherche en éducation. Certains clusters partagent des références dont les publications sont très connectées et portent directement sur l'enseignement et l'apprentissage (e.g. le cluster *Learning*) alors que d'autres clusters partagent des références de manière moins connectée et s'intéressent à des sujets qui se trouvent à l'intersection des recherches autour du fait éducatif et d'autres champs d'investigation (e.g. le cluster *Sociology of Education*). En résumé, ces clusters représentent des champs de recherche en éducation qui sont organisés de manière différente autour de sujets variés et qui partagent des bases de connaissances différentes, incarnées dans des références partagées. Nous proposons que des étudiants, chercheurs, et praticiens utilisent nos résultats pour explorer ces bases de données sur le fait éducatif, construites de manière collective.

Mots clefs: recherche en éducation, couplage bibliographique, scientométrie

Introduction

La scientométrie, science des publications scientifiques, est un champ de recherche qui utilise diverses métriques afin de comprendre le développement des sciences et des pratiques scientifiques (Mingers & Leydesdorff, 2015). Dans cet article, nous utilisons des méthodes scientométriques afin de comprendre la carte globale de la recherche en éducation — dans le contexte de la base de données internationale *Scopus*. En particulier, notre approche est centrée sur une étude de citations. Etant donné que les publications académiques sont connectées au travers de leurs citations, celles-ci peuvent être vues comme des réactions à une contribution faite au sein d'une communauté de connaissance, dans une approche similaire à l'étude du même phénomène lors de débats au sein d'une interaction en petit groupe (Suthers, Dwyer, Medina, & Vatrappu, 2010). Quand une référence est citée, elle est ancrée dans la construction collective de connaissances par sa connexion à d'autres travaux dans une base de connaissances partagées, similaire au processus de *grounding*, qui a lieu lors d'une conversation (Clark, 1996). Même une contribution individuelle remarquable ne devient valable qu'intégrée avec le reste des travaux d'une communauté (Rosé & Lund, 2013). A travers des analyses du champ des recherches du comportement humain, Longino (2013) a conclu que l'ensemble des références citées par un article donné est très spécifique à certaines disciplines et aux communautés de lecteurs auxquelles la recherche est destinée et que très peu de citations sont communes à des chercheurs utilisant différentes approches. En somme, « Dites-moi qui vous citez, et je vous dirai à quel champ vous appartenez » est une vérité générale qui est au cœur de la scientométrie. Dans cet article, nous examinons des patterns de citations dans la recherche autour du fait éducatif afin de révéler la nature des sous-ensembles de champs de recherche en éducation, pour la période étudiée.

Recherches antérieures dans le champ de l'éducation

Une grande partie des recherches en éducation qui ont employé des méthodes scientométriques a visé des champs restreints pour des raisons diverses. Par exemple, (Kirby, Hoadley, & Carr-Chellman, 2005) ont exploré la relation entre les *Learning Sciences* et le *Instructional Systems Design* à travers une analyse de citations dans six revues internationales, trois dans chaque champ. Leur analyse montre que le nombre de citations trans-champs était basse, bien que son évolution était croissante et que ces citations étaient faites par un petit nombre de chercheurs prééminents, dans chaque communauté. Sur la base d'une analyse qualitative, les auteurs ont argumenté qu'une combinaison des points forts de chaque communauté — la cognition en contexte (*Learning Sciences*) d'un côté et la conception (*Instructional Systems Design*) de l'autre, pourraient favoriser la création et l'insertion des environnements informatiques pour l'apprentissage humain dans des contextes particuliers (op. cit., p. 46).

Kienle et Wessner (2006) ont exploré la croissance du champ *Computer Supported Collaborative Learning* (CSCL) par l'analyse de dix ans de conférences CSCL (1995-2005). Ils ont employé une analyse mixte faite d'une part, d'une approche quantitative (analyses de citations et de co-auteurs, couplée avec une analyse des participants et comités de programmes de ces conférences) et d'autre part, d'une approche qualitative de questionnaires ouverts. Les résultats ont montré qu'un petit noyau de personnes avait participé de manière continue à la communauté, bien que la participation se soit internationalisée à ce moment-là et que la connectivité internationale ait été aussi en croissance.

Une autre étude par Tang, Tsai, et Lin (2014) a analysé la structure thématique du champ CSCL en utilisant plus de mille articles dans des revues internationales et conférences renommées, indexées dans le *Web of Science* entre 2006 et 2013. Les résultats ont permis d'identifier plusieurs courants de recherche : 1) la représentation, le discours et des patterns, 2) les facteurs influençant les CSCL, 3) l'intervention et la comparaison, 4) le raisonnement critique, 5) le processus de la construction sociale, et 6) la conception et la modélisation des CSCL. De plus, 56 des 66 papiers considérés comme fondamentaux, étant donné leur niveau de citations se trouvaient dans les trois premiers courants de recherche. Ces courants peuvent alors être considérés comme les plus représentatifs du champ.

Chacun de ces trois articles (Kienle & Wessner, 2006; Kirby et al., 2005; Tang et al., 2014) se focalisait sur un champ précis touchant à la recherche en éducation : le *Learning Sciences* et *Instructional Systems Design* d'un côté et le *Computer Supported Collaborative Learning* de l'autre. Ces articles portaient sur les bases de connaissances de ces communautés, ou bien sur les chercheurs qui les ont créées. Dans notre article, nos objectifs sont similaires, mais nous tendons vers une description plus générale de la recherche en éducation — dans les limites de la base de données *Scopus*. Nous nous intéressons à la caractérisation des clusters bibliographiques des articles de recherche en éducation et dans quelle mesure chacun de ces clusters correspond à des champs particuliers de la recherche.

Méthodes

Dans cette section, nous décrivons tout d'abord la sélection de notre base de données ainsi que l'extraction des données. En deuxième lieu, nous décrivons le concept du couplage bibliographique (CB) et comment nous avons construit le réseau des clusters concernant la recherche autour du fait éducatif.

Sélection d'une base de données et extraction de ces données

Il existe plusieurs bases de données qui indexent des publications académiques. Chacune d'entre elles offre une couverture étendue d'une sélection de publications. Nous avons choisi d'utiliser *Scopus* qui a une couverture similaire au *Web of Science*, mais plus étendue concernant les revues non-anglophones. Les sources de publications dans *Scopus* sont

catégorisées sous quatre grands sujets (*Life Sciences, Physical Sciences, Health Sciences* et *Social Sciences & Humanities*) et ceux-ci sont encore divisés en 27 sujets plus précis. Comme l'interface web de *Scopus* n'offre pas un sujet correspondant à l'éducation, nous nous sommes appuyés sur la liste de publications en éducation fournie par l'AERES : l'Agence d'Evaluation de la Recherche et de l'Enseignement Supérieure (AERES, 2014). Cette liste a été basée à son tour sur la liste issue du projet EERQI: *European Educational Research Quality Indicator* (<http://www.eerqi.eu/>). Il est clair que le pouvoir de décider ce qui se qualifie comme recherche en éducation réside dans les organismes responsables de la constitution de ces listes.

Afin de constituer le CORPUS EDUCATION étudié dans cet article, nous avons compilé les caractérisations les plus complètes des publications: auteurs, nom de la source de publication, année de publication, titre de publication, mots clés (choisis par les auteurs ou donnés par *Scopus*), et les listes des références incluses dans les publications. Dans ce travail initial, nous avons uniquement recensé des publications de 2000-2004 puisqu'après cette période, il y a eu une grande augmentation dans le nombre de revues publiées et acceptées dans la base de données *Scopus*. Des travaux similaires sont en cours pour des publications plus récentes ainsi que sur l'évolution de la recherche en éducation. A partir des listes de références d'une publication donnée, nous avons extrait des auteurs, l'année de publication et le titre de chaque référence. Les références n'ont pas toujours été formatées de manière consistante par *Scopus* (e.g. utilisation d'abréviations différentes, informations manquantes,...). En tout, environ 2% des références étaient formatées de manière incorrecte. Dans sa totalité le CORPUS EDUCATION contient 36,715 enregistrements bibliographiques. Les publications sont écrites majoritairement en anglais (94%) par des auteurs des Etats-Unis (44%), suivi par des auteurs du Royaume Uni (14%) et d'Australie (5%). La carte EducMap ainsi qu'une description des métadonnées agrégées du corpus est disponible en ligne avec un accès libre¹. Plusieurs collaborations sont en cours pour étendre les analyses jusqu'en 2014.

Les deux domaines qui ont le plus contribué aux publications de ce CORPUS EDUCATION portent sur l'éducation en chimie (2532 items) et en médecine (1162), issues du *Journal of Chemical Education* et de *Medical Education*. Les trois sources suivantes en terme de contributions au corpus, sont des magazines professionnels pour des praticiens : *Educational Leadership* (727), *Phi Delta Kappan* (607), et une source sur la recherche du développement des enfants : *Child Development* (609). Plusieurs autres sources se concentrent sur des domaines d'éducation en sciences dites « exactes » : *International Journal of Engineering Education* (492), *International Journal of Science Education* (393), *IEEE Transaction on Education* (373), et *Biochemistry and Molecular Biology Education* (371). Un autre type de source étudie l'enseignement de différents domaines : *Medical Teacher* (590), *The Reading Teacher* (397), et *Teaching of Psychology* (388). Enfin deux sources montrent à nouveau la présence de la psychologie : *Journal of Educational Psychology* (337) et la préoccupation du développement et de la gestion des environnements informatiques *Educational Technology and Society* (343).

Couplage bibliographique et construction du réseau de clusters

Afin de déterminer comment différentes publications sont liées à travers de références communes, nous avons comparé de manière systématique la liste des références de toutes les paires de publications et identifié leurs références partagées. Deux publications sont liées si elles partagent au moins deux références, ce qui donne lieu à un réseau de publications connectées. Ce réseau est représenté de manière schématique en Figure 1a où les nœuds (cercles) représentent des publications individuelles éventuellement connectées par des liens. Un lien plus épais indique un plus grand nombre de références partagées. un 'poids' ω_{ij} est

¹ <http://le.ens-lyon.fr/ressources/educmap>

donné aux liens en fonction de la similarité des références des deux articles. Cette similarité est mesurée via le cosinus de Kessler

$$\omega_{ij} = \frac{|R_i \cap R_j|}{\sqrt{|R_i||R_j|}}$$

où $|R_i|$ est le nombre de références de l'article i et $|R_i \cap R_j|$ le nombre de références partagées par R_i et R_j . Par définition, la similarité est égale à zéro quand deux articles ne partagent pas de références et est égale à 1 quand leurs ensembles de références sont identiques.

Les clusters sont alors détectés automatiquement par un algorithme informatique qui construit des cartes où la densité de liens est maximale dans les clusters et minimale entre clusters. Cet algorithme utilise une méthode de maximisation de modularité (Newman & Girvan, 2004) basée sur le *fast Louvain Algorithm* (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). La modularité quantifie la possibilité de diviser le réseau en clusters. Il y a plusieurs techniques disponibles pour partitionner des nœuds d'un graphe en 'communautés' pertinentes (cf. Fortunato, 2010 pour une revue). Grâce à sa simplicité et à sa facilité de programmation, la modularité est de moyen le plus populaire, même si ses résultats doivent parfois être interprétés avec prudence (Good, de Montjoye, & Clauset, 2010). Dans d'autres recherches sur des réseaux bibliographiques similaires (Grauwin & Jensen, 2011; Grauwin et al., 2012), nous avons montré que les clusters obtenus par la maximisation de la modularité représentent bien la structure scientifique des champs concernés.

La Figure 1 montre schématiquement comment les clusters se forment petit à petit ; les couleurs servent à distinguer des clusters qui émergent par la comparaison de publications, deux à deux. Notez que des publications qui appartiennent au même cluster (e.g. nœud 1 et 5 ou nœuds 12 et 15) ne sont pas toujours liés de manière directe. Notez également que des publications qui appartiennent à différents clusters peuvent aussi partager des liens. La méthode Louvain détecte des clusters de nœuds tels que le nombre de connexions externes soit le plus petit possible. Dans les réseaux étudiés dans cet article, plus de 70% de liens d'une publication donnée sont avec des publications appartenant au même cluster.

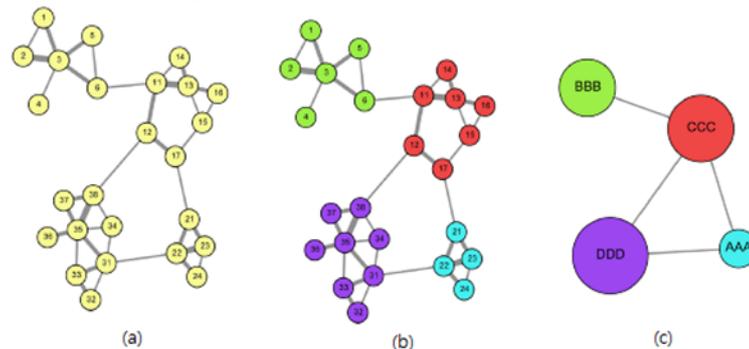


Figure 1. Processus pour détecter des clusters

Résultats: Champs en recherche autour du fait éducatif entre 2000-2004

Cette section illustre les résultats et l'interprétation de nos analyses scientométriques de la base de données *Scopus* de 2000 à 2004. Tout d'abord, nous présentons les clusters principaux résultant de notre analyse de couplage bibliométrique sur le CORPUS EDUCATION. Deuxièmement, nous illustrons la nature des références les plus citées de trois des clusters. Puis, nous examinons la connectivité de ces clusters, distinguant cinq clusters dont les publications sont très connectées et dont l'éducation est le sujet central. Les publications des autres clusters sont moins connectées et s'intéressent ainsi aussi à d'autres sujets, en plus des sujets autour du fait éducatif.

Clusters principaux résultant de l'analyse du couplage bibliographique

Nous avons appliqué l'algorithme de détection des clusters aux 36,715 enregistrements du CORPUS EDUCATION. 22,058 entre eux partageaient au moins deux références avec d'autres données dans le corpus et sont donc inclus dans le réseau final de clusters. La table 1 représente les 18 clusters, qui en résultent, chacun incluant un minimum de 400 articles. Afin de leur donner des étiquettes pertinentes, nous avons examiné les références les plus partagées ainsi que les mots des titres, ainsi que les mots clefs, des articles les plus utilisés. Ces étiquettes illustrent le fait que chaque cluster se focalise sur différents aspects de l'éducation, et s'appuie sur un ensemble différent de références partagées.

Table 1. Les dix-huit clusters bibliographiques principaux de la recherche en éducation (Scopus 2000-04)

Clusters organisés par taille	N	Clusters organisés par taille	N
<i>Learning</i>	1,883	<i>Cognitive Studies of Learning</i>	790
<i>Educational Equality</i>	1,800	<i>Assessment & Evaluation</i>	733
<i>Sociology of Education</i>	1,715	<i>Math Education</i>	685
<i>Child Behavioral Development</i>	1,534	<i>Language Teaching Methods</i>	675
<i>Motivation</i>	1,514	<i>Developmental Disabilities</i>	667
<i>Science Education</i>	1,370	<i>Measurement</i>	648
<i>Higher Education</i>	1,207	<i>Cooperative Learning</i>	554
<i>Reading Education</i>	1,140	<i>Civic Education</i>	417
<i>Teacher Training</i>	799	<i>Child Cognitive Development</i>	415

Certains clusters sont centrés sur l'enseignement d'un sujet précis comme les mathématiques, la science, ou une langue. D'autres étudient des mécanismes sous-jacents à l'enseignement-apprentissage comme la motivation et le développement cognitif des enfants. D'autres encore travaillent sur des sujets de société plus larges comme l'égalité et l'impact de l'éducation. La diversité des sujets de recherche couverte par les différents clusters explique en partie pourquoi la recherche en éducation s'appuie sur des ensembles multiples de références. Bien que toute la recherche en éducation ne soit pas représentée ici, l'ensemble de ces clusters illustre le fait que le domaine ne porte pas seulement sur les contenus des enseignements et sur les mécanismes fondamentaux de l'apprentissage et du développement, mais également sur des sujets sociaux, historiques, politiques, économiques et techniques.

La nature des références les plus citées dans trois clusters

Nous avons examiné la nature des références les plus citées dans trois des clusters (*Learning*, *Assessment & Evaluation*, et *Measurement*). Pour cela, nous avons codé leurs vingt références les plus citées, que nous appelons références noyau (cf. la table 2 pour les cinq premières) selon leurs objets 1) théoriques, 2) méthodologiques, 3) empiriques, 4) basés sur des connaissances disciplinaires (étiqueté *domaine*), ou 5) reflets d'un consensus d'experts et de praticiens dans les communautés en question (par exemple, des standards d'évaluation pour les programmes scolaires). Le jugement de codage était basé sur notre connaissance de la littérature en éducation et sur le titre et les contenus des articles.

Table 2. Les cinq références les plus citées des trois clusters choisis, les citations complètes sont dans nos références

<i>Learning</i>	<i>Assessment & Evaluation</i>	<i>Measurement</i>
Lave & Wenger, 1991	Marton & Säljö, 1976	Wilkinson, 1999
Vygotsky, 1978	Ramsden, 1992	Lord, 1980
Brown, Collins, & Duguid, 1989	Entwistle & Ramsden, 1983	<i>Standards for educational and psychological testing</i> , 1999
Wenger, 1998	<i>Higher education in the learning society</i> , 1997	Lord, 1968

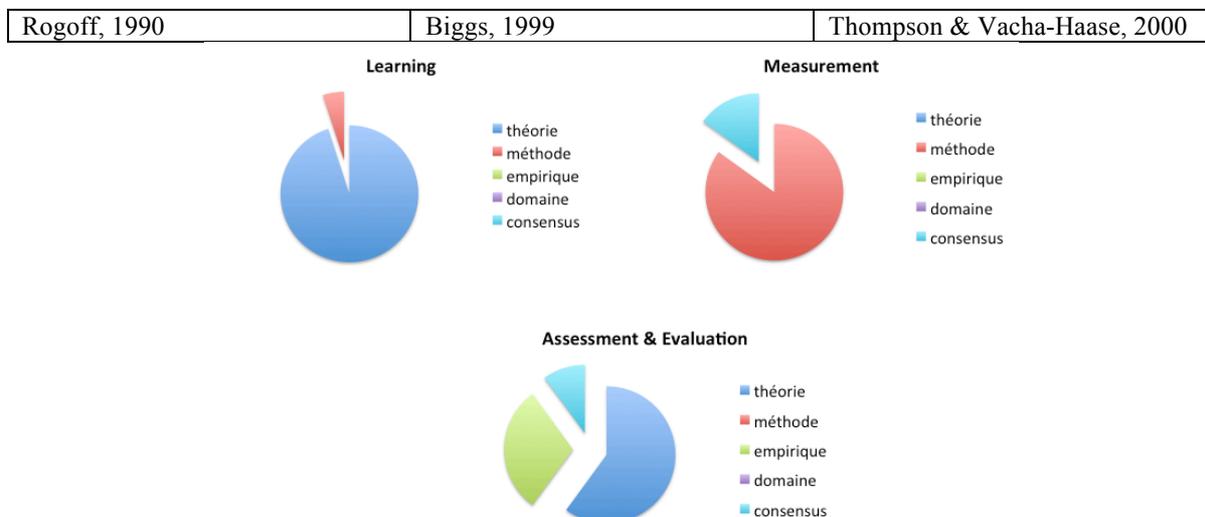


Figure 2. Une comparaison de la nature des vingt plus cités références de trois clusters

L'application de la méthode des juges a montré un résultat élevé d'accords entre codeurs (92.5%). Les quelques avis divergents portaient sur une hésitation quant à la caractérisation prédominante : une contribution théorique d'un article versus une contribution empirique.

La Figure 2 montre que le type de références qui connectent les articles dans le cluster *Learning* est très majoritairement théorique. A cette époque, il s'agissait de théories socio-culturelle, comme l'apprentissage situé de Lave and Wenger (1991) et l'interaction entre l'apprentissage et le développement de Vygotsky (1978). Par contre, dans le cluster *Measurement*, les références sur les méthodes (e.g. la théorie de la réponse item) et le consensus (e.g. des standards pour les tests psychologiques et éducationnels) jouent un rôle primordial. Enfin, dans le cluster *Assessment & Evaluation*, la plupart des références sont de nature théorique sur l'apprentissage, suivie par des travaux empiriques et des documents de consensus. Les vingt premières références les plus citées (références de noyau) de chacun des clusters ne se chevauchent que rarement. Par exemple, Vygotsky (1978) et Lave & Wenger (1991) sont les deux références les plus citées à travers tous les clusters, mais ne sont présentes dans les vingt références les plus citées, que dans un des clusters (i.e. *Learning*).

La connectivité des clusters

Dans cette section, nous illustrons comment les différents clusters sont liés entre eux. L'épaisseur d'un lien entre deux clusters dans la Figure 3 est proportionnelle au nombre de références qu'ils partagent. Nous parlons du poids (P) entre les publications des deux clusters. Notez que ce lien *entre* clusters émerge des références partagées qui ne sont pas dans les références noyau de chaque cluster (i.e. les vingt premières citées).

Si les poids de tous les liens entre paires de clusters sont additionnés, nous obtenons une mesure de la connectivité totale d'un cluster. Dans le vocabulaire de la théorie des réseaux, le terme « centralité » est utilisé pour indiquer le degré de connectivité générale du cluster au sein du réseau (cf. Table 3). Ce terme n'implique pas que ces travaux soient centraux pour la recherche en éducation ; ceci est une question à la quelle il faut répondre avec d'autres indicateurs. Les clusters avec une valeur P élevée partagent plus de références avec d'autres clusters que des clusters avec une valeur P basse, qui eux utilisent des références plus spécialisées. L'algorithme qui est utilisé pour visualiser les clusters en réseau tend à positionner des clusters qui sont connectés de manière forte au centre de la carte. Les clusters qui sont moins connectés avec d'autres se trouvent à la périphérie de la carte (cf. Figure 3).

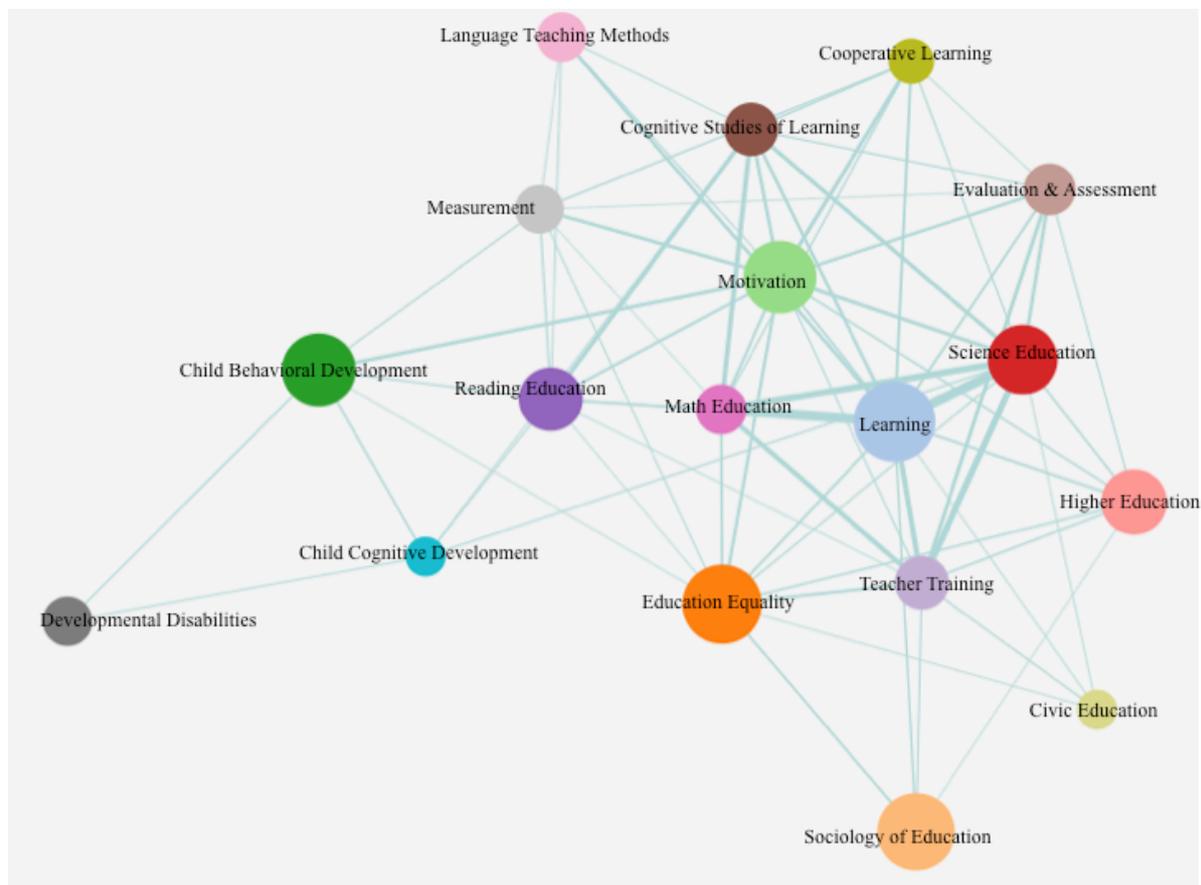


Figure 3. Une carte montrant les 18 différents clusters de recherche en éducation entre 2000 et 2004, a été construite algorithmiquement par une analyse de couplage bibliométrique de notre base de données. L'épaisseur des liens est proportionnelle à l'intensité de références partagées

Les clusters diffèrent également en terme de modularité interne Q_i , une caractérisation de leur hétérogénéité interne (cf. Table 3). Il s'agit d'une mesure de la partition de chaque cluster i en sous-clusters. Des valeurs basses de Q_i correspondent à des clusters contenant des publications qui sont liées de manière homogène (e.g. *Learning*). Cela signifie que la plupart des publications partagent des références appartenant au noyau de référence (cf. Figure 4a). Par exemple, les deux références les plus citées du cluster *Learning* sont citées par une fraction élevée de ses publications (respectivement 19% pour Lave & Wenger, 1991 et 12% pour Vygotsky, 1978). Des valeurs élevées de Q_i correspondent à des clusters contenant des publications qui sont liées de manière hétérogène. Par exemple, le cluster *Sociology of Education* (cf. Figure 4b) pourrait être composé de plusieurs sous-clusters liés par des ensembles relativement distincts de références. Cela résulte en une fréquence nettement plus basse des deux références les plus citées pour ce cluster (5% pour Gewirtz, Ball, & Bowe, 1995 et 4% pour Olivier, 1996).

Les cinq clusters les plus centraux, c'est à dire étant les plus susceptibles de partager des références avec d'autres clusters sont *Learning*, *Motivation*, *Science Education*, *Math Education*, et *Teacher Training*). Ces clusters apparaissent comme se focalisant directement sur l'apprentissage et l'enseignement. Cette focalisation peut se réaliser en terme d'approches théoriques à l'éducation (e.g. visions socioculturelles), d'attitudes et d'identité des enseignants et des élèves ou en terme de pratiques en classe. L'angle d'étude peut être également plus spécifique en terme de contenus, par exemple, sur des pratiques d'enseignement et d'apprentissage en mathématique et en science.

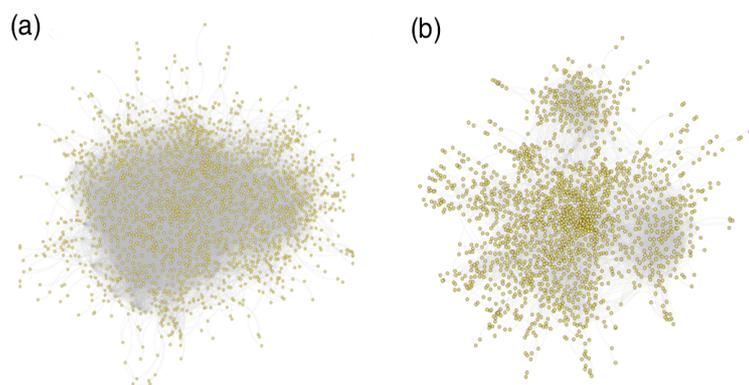


Figure 4. Une illustration de la différence entre des clusters qui sont liés de manière homogène (a) et de manière hétérogène (b)

Table 3. Clusters sont catégorisé du plus haut vers le plus bas en terme de P, où la somme de tout P égale à 100.

Clusters	N	P (%)	Qi		N	P (%)	Qi
<i>Learning</i>	1,883	11.2	0.43	<i>Measurement</i>	648	5.3	0.61
<i>Motivation</i>	1,514	10.1	0.47	<i>Higher Education</i>	1,207	4.3	0.68
<i>Science Education</i>	1,370	10.1	0.48	<i>Cooperative Learning</i>	554	4.2	0.79
<i>Math Education</i>	685	8.5	0.36	<i>Child Behav. Development</i>	1,534	3.4	0.71
<i>Teacher Training</i>	799	8.2	0.64	<i>Language Teaching Methods</i>	675	2.8	0.57
<i>Cognitive Studies of Learning</i>	790	6.8	0.59	<i>Sociology of Education</i>	1,715	2.5	0.71
<i>Assessment & Evaluation</i>	733	5.7	0.59	<i>Child Cognitive Development</i>	415	2.5	0.64
<i>Educational Equality</i>	1,800	5.4	0.67	<i>Civic Education</i>	417	2.1	0.79
<i>Reading Education</i>	1,140	5.3	0.43	<i>Developmental Disabilities</i>	667	1.5	0.82

Par contre, les clusters connectés de manière moins étroite (i.e. *Language Teaching Methods*, *Sociology of Education*, *Child Cognitive Development*, *Civic Education*, and *Developmental Disabilities*) peuvent être subdivisés en de multiples sous-clusters qui reflètent une granularité plus fine. Dans le cluster, *Developmental Disabilities*, par exemple, le plupart des sous-clusters que l'on peut construire sont liés aux domaines médicaux ou à la psychologie, ce qui est démontré par leur usage élevé du *Diagnostic and Statistical Manual of Mental Disorders*, publié de manière collective par le *American Psychiatric Association*. De manière similaire, l'éducation n'est qu'un exemple d'un phénomène sociologique dans le cluster *Sociology of Education*, qui traite également d'autres sujets sociologiques comme la globalisation ou le colonialisme. Le cluster *Language Teaching Methods* consiste également en de multiples sous-clusters qui portent non seulement sur des méthodes d'enseignement de langue, mais également sur des sujets comme la politique linguistique ou la complexité syntaxique dans la communication asynchrone versus la communication synchrone. Dans ce dernier cas, comprendre de sujets aussi divers est essentiel dans la conception et l'implémentation de méthodes d'enseignement en langue et pointe vers la nécessaire pluridisciplinarité de la recherche en éducation. Selon nos analyses, la façon dont différents sous-champs en éducation s'appuient sur les références d'autres sous-champs (en linguistique, sociologie, psychologie, etc.) diffère selon les clusters. Certains s'appuient plutôt sur des références dans d'autres sous-champs de l'éducation, d'autres s'appuient sur des références extérieures au champ de l'éducation.

De plus, il y a une forte corrélation ($R^2=0.58$) entre le degré pondéré (P) des clusters et leur modularité interne (Q_i): les clusters homogènes (Q_i basse) tendent à être des clusters très

connectés (P élevé). Nous interprétons cette corrélation comme une conséquence du filtrage de la littérature scientifique par notre liste de revues en éducation, qui — encore une fois — définit de manière implicite une certaine vision de la recherche en éducation. Pour certains sous-champs, cela permet d'avoir accès à leurs publications principales. Notre analyse peut donc récupérer la cohérence de leur « noyau de connaissance » et ces sous-champs apparaîtront comme homogène (quantifiés par une Q_i basse) et très connectés.

Par contre, pour d'autres sous-champs, l'intersection de leurs publications principales avec notre définition de la recherche en éducation peut, peut-être, briser leur cohérence (supposée dans ce cas). Il s'agit là des clusters *Language Teaching Methods*, *Sociology of Education*, *Child Cognitive Development*, *Civic Education*, et *Developmental Disabilities*). Ces sous-champs apparaissent comme la juxtaposition de clusters connectés de manière moins étroite entre eux. Certains sont connectés à des publications en éducation, mais la plupart ne le sont pas. Cela est quantifié à la fois par une Q_i élevé (sous-clusters connectés de manière lâche) et par un poids bas (P) où la plupart de publications ne sont pas connectées au domaine de l'éducation.

Synthèse, conclusions et perspectives

Nous avons présenté une analyse scientométrique de publications en éducation de la période 2000-2004, utilisant la base de données *Scopus*. La période a été choisie pour sa stabilité statistique puisque juste après, il y a eu une forte augmentation à la fois du nombre de revues publiées et de celles prises en compte par *Scopus*. Dans la première section, nous avons construit une carte de la recherche globale en éducation qui consiste en 18 clusters ou champs d'investigation. Chacun de ces clusters représente un ensemble de publications qui sont liées par des références partagées. La nature des références varie selon les clusters. Certains partagent une majorité de références théoriques, d'autres, une majorité de références méthodologiques et d'autres encore un mélange de types de références. Une caractéristique notable est la haute visibilité des documents de consensus comme les *Standards for educational and psychological testing*. Ce type de publication ne joue pas de rôle prédominant dans d'autres champs de recherche. Par exemple, aucun document de consensus n'apparaît dans le noyau des références les plus citées des clusters de certaines disciplines (Grauwin, et al., 2012). Cela dit, il est évident que les méthodes d'enseignement et les objectifs d'apprentissage sont des sujets d'interrogation déterminés non seulement par des principes scientifiques, mais également par ce que la société valorise comme connaissances. Cette valorisation est fortement influencée par des contextes historiques et culturels. Les tensions qui émergent des perspectives nécessairement diverses sur ces sujets requièrent des efforts de construction de consensus, reflétés dans les documents de cette nature. De tels consensus sont informés par la recherche, mais fonctionnent également comme un guide pour la recherche en éducation. Il y a donc un effet de causalité réciproque.

Les clusters diffèrent également en terme de leur connectivité envers d'autres clusters. Dans notre analyse, certains clusters (e.g. *Learning*, *Motivation*, *Science Education*, *Math Education*, et *Teacher Training*) sont plus connectés que d'autres. Ces clusters et leurs références associées, qu'elles portent sur des théories ou sur des méthodes, jouent un rôle plus central dans l'organisation et la cohérence de la recherche en éducation, du moins dans la vision donnée par la base de données internationales *Scopus*. Dans le cas du cluster *Learning*, le cluster avec la connectivité la plus haute, les références théoriques de l'approche socioculturelle ont joué un rôle prédominant et ont inclus les deux références les plus citées sur tout le CORPUS EDUCATION : (Vygotsky, 1978; Lave & Wenger, 1991). Ces deux références étaient également parmi celles qui liaient tous les clusters entre eux. Nous pouvons alors faire l'hypothèse que la théorie socioculturelle représente un paradigme principal pour cette période — dans la limite de notre base de données. La connectivité des clusters avec d'autres

clusters était hautement corrélée avec leur hétérogénéité interne et les clusters les moins connectés entre eux consistent en de multiples sous-clusters. Certains de ces sous-clusters s'appuient sur des références en dehors de l'éducation. Ce constat démontre la nature pluridisciplinaire de la recherche en éducation et la nécessité de se pencher sur les possibilités d'intégration de la recherche en éducation avec le reste des sciences humaines et sociales notamment la sociologie, la psychologie, les sciences du langage, les sciences politiques, l'histoire, etc.

Malheureusement, il n'y a pas actuellement de bases de données en français comparables à *Scopus*, c'est à dire incluant des métadonnées, qui permettraient une analyse semblable de la recherche francophone. Dans nos perspectives de recherche, nous comptons travailler tout de même au décloisonnement des disciplines de manière qualitative à la fois dans un contexte francophone et au niveau international (Lund, 2016). Des travaux sont en cours dans ce sens au sein du Laboratoire de l'Education, financés conjointement par le CNRS et par l'Ecole Normale Supérieure de Lyon, et impliquant la structure fédérative RELYS (Recherches en Education Lyon-Saint-Etienne).

Références

- American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC, American Psychological Association.
- Biggs, J. (1999). *Teaching for Quality Learning at University: What the Student Does*. Buckingham: Open University Press.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks, 6. doi:10.1088/1742-5468/2008/10/P10008
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 32–42.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Entwistle, N. and Ramsden, P. (1983) *Understanding Student Learning*, London: Croom Helm.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3), 75–174.
- Gewirtz, S., Ball, S. & Bowe, R. (1995) *Markets, Choice and Equity in Education*. Buckingham: Open University Press.
- Good, B. H., de Montjoye, Y. A., & Clauset, A. (2010). Performance of modularity maximization in practical context. *Physical Review E*, 81(4), 046106.
- Grauwin, S., Beslon, G., Fleury, É., Franceschelli, S., Robardet, C., Rouquier, J.-B., & Jensen, P. (2012). Complex systems science: Dreams of universality, interdisciplinarity reality. *Journal of the American Society for Information Science and Technology*, 63(7), 1327–1338. doi:10.1002/asi.22644
- Grauwin, S., & Jensen, P. (2011). Mapping scientific institutions. *Scientometrics*, 89(May), 943–954. doi:10.1007/s11192-011-0482-y
- Great Britain. National Committee of Inquiry into Higher Education & Dearing, Ron (Sir), 1930- (1997). Higher education in the learning society. The Committee, [Leeds].
- Kessler, M. M. (1963). Bibliographic coupling between scientific articles. *American Documentation*, 24, 123–131.
- Kienle, A., & Wessner, M. (2006). The CSCL community in its first decade: development, continuity, connectivity. *International Journal of Computer-Supported Collaborative Learning*, 1(1), 9–33. doi:10.1007/s11412-006-6843-5
- Kirby, J. A, Hoadley, C. M., & Carr-Chellman, A. a. (2005). Instructional systems design and the learning sciences: A citation analysis. *Educational Technology Research & Development*, 53(1), 37–47. doi:10.1007/BF02504856

- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Longino, H. E. (2013). *Studying human behavior: How scientists investigate aggression and sexuality*. Chicago: University of Chicago Press.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Mahwah, NJ: Erlbaum.
- Lund, K. (2016). Décloisonner la recherche en éducation pour offrir des nouvelles opportunités scientifiques. Dans P. Bourdelais (Ed.) *La lettre de l'InSHS*. p. 25-26. http://www.cnrs.fr/inshs/Lettres-information-INSHS/lettre_infoinshs40hd.pdf
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: outcome and process. *British Journal of Educational Psychology* 46(1), 4–11.
- Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research*, 1–19.
- Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 026113.
- Oliver, M. (1996). *Understanding disability: from theory to practice*. Basingstoke: Macmillan.
- Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Rosé, C. P., & Lund, K. (2013). Methodological pathways for avoiding pitfalls in multivocality. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 613–637). New York: Springer.
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrappu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction, 5–42. doi:10.1007/s11412-009-9081-9
- Tang, K.-Y., Tsai, C.-C., & Lin, T.-C. (2014). Contemporary intellectual structure of CSCL research (2006–2013): a co-citation network analysis with an education focus. *International Journal of Computer-Supported Collaborative Learning*, 9(3), 335–363. doi:10.1007/s11412-014-9196-5
- Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, 60, 174-195.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, MA: Harvard University Press.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.

Remerciements

Ces recherches ont été financé par le (PEPS) Université de Lyon – CNRS 2015 et par Aslan (ANR-10-LABX-0081) de l'Université de Lyon, au sein des *Investissements d'Avenir* (ANR-11-IDEX-0007). Heisawn Jeong était professeure invitée à l'ENS Lyon durant 2014 et son séjour était financé à la fois par ENS-Lyon et par l'université de Hallym en Corée du Sud (HRF-201408-004).