

Mettre les pendules à l'heure : Preuve scientifique, preuve statistique, preuve culturelle

Jean-Marie Van der Maren, professeur honoraire, Université de Montréal
Clermont Gauthier, professeur émérite, Université Laval
Steve Bissonnette, professeur titulaire, TELUQ

Les débats acrimonieux, et plus politiques que scientifiques, que certains universitaires ont tenus avant que la loi 23 ne soit votée à l'Assemblée nationale du Québec et que l'Institut National d'Excellence en Éducation (INNE) n'ait été institué, sont terminés. Ils ont été l'occasion de lire des déformations parfois explicites, souvent implicites, du sens que certains termes techniques ont en recherche et des désinformations quant à l'histoire d'une discipline que l'on nomme (parfois abusivement) « sciences » de l'éducation tout en confondant éducation et pédagogie. L'objectif de ce texte est de corriger certaines erreurs, certaines confusions et certains oublis, dans lesquels certaines publications ont glissé sans doute à cause de leurs visées plus polémiques dans les débats de l'époque.¹

Précisons ici que par pédagogie nous entendons la discipline qui s'intéresse à l'instruction, aux premières et fondamentales missions de l'École : faire apprendre aux jeunes à lire et comprendre ce qu'ils lisent, à écrire de telles manières que les lecteurs éventuels les comprennent, à compter sans se tromper, à comprendre et à parler une langue seconde, à connaître les grands moments de leur histoire nationale et de l'histoire mondiale, à savoir situer géographiquement les autres humains avec leur économie et leur culture. Par ailleurs, nous entendons par « éducation » tout ce qui peut entrer dans ce qui était, il n'y a même pas un siècle, du rôle des parents, des organisations religieuses (églises, synagogues et mosquées) et des organisations politiques : la sexualité, le civisme, le sens de la justice et du respect, le bien commun, etc., De plus, lorsqu'on adopte une définition de la finalité de l'école holistique comme étant la réussite éducative des élèves, « c'est-à-dire l'expression de leur plein potentiel dans les différents aspects du développement global »² (Altet, 2002, repris par Pulido et Allaire, 2023), on ne se trouve plus en pédagogie mais en éducation au sens large et cela déborde amplement la tâche des enseignants.³

1. De l'origine de la recherche en pédagogie (éducation)

Ainsi, une première erreur lue et relue (dont une référence à Pulido et Allaire, 2023) situe en médecine l'origine de la recherche évaluative recourant à des plans expérimentaux et quasi expérimentaux. Ce qui permet à plus d'un de vouloir décoloniser ou émanciper l'éducation et la recherche en éducation de la médecine, plus spécifiquement de la pratique médicale perçue, fantasmatiquement, comme une victime du dictat de la recherche

¹ Sans le viser particulièrement, notre texte répond en quelque sorte au chapitre de Loïc Pulido et Stéphane Allaire, tous deux de l'UQAC, intitulé : Proposition pour définir et mobiliser les « pratiques efficaces reconnues par la recherche » et inséré le 10 novembre 2023 dans Laferrière T. & Coll. 2023.

² Marguerite Altet impliquait jusqu'à un psychologue psychanalyste, un linguiste, un épistémologue et un sociologue dans son analyse plurielle. On est loin du regard pédagogique par le didacticien de son équipe.

³ Débordement de la tâche dont se sont largement plaint les enseignants lors négociations (2023-2024) des dernières conventions collectives.

scientifique et des données probantes ne laissant guère d'initiatives aux cliniciens. Or, c'est bien une certaine liberté d'initiative dans l'action pédagogique que les syndicats revendiquent, ce qui leur permet de rejeter violemment un parallèle tout aussi fantasmagique que pourraient être tentés de faire tant les politiques que les chercheurs « positivistes » des sciences de l'éducation.

Or, l'origine des recherches expérimentales et évaluatives en éducation se trouve, non pas en médecine, mais à la fin du XIXe siècle (entre 1890 et 1912) lorsque les discours pédagogiques prémodernes vont progressivement céder le pas, dans les instituts universitaires, à une perspective moderne, scientifique, lorsque la rationalité scientifique remplace la prescription religieuse ou, plus rarement, une morale humaniste. Les tenants de la pédologie initiée par Oscar Chrisman (1893), professeur de pédologie et de psychologie à l'Université de l'Ohio (il avait fait son doctorat à Iena en Allemagne), veulent fonder une discipline pratique objective basée sur la connaissance scientifique du développement général des enfants et sur l'évaluation scientifique (expérimentale) de l'efficacité des dispositifs d'enseignement. Une illustration en est le « Laboratoire de Pédologie » dirigé par Schuyten à Anvers (1899) et le premier Congrès de Pédologie qui se tient aussi à Anvers en 1912. Le modèle est alors emprunté aux sciences naturelles qui prônent que la connaissance scientifique découle de manière inductive de la recherche expérimentale.

Un des outils privilégiés de la recherche expérimentale et quasi expérimentale, après les plans de recherche, est l'utilisation des statistiques, d'abord descriptive puis inférentielle. Cette utilisation est nécessaire afin de comparer les résultats fournis par plusieurs sujets sinon par plusieurs groupes de sujets et pas seulement les résultats d'un ou de quelques sujets avant et après l'intervention à propos de laquelle on se pose des questions, que ce soit d'efficacité ou même d'efficience. Notons de plus, en ce qui concerne les statistiques et les comparaisons entre groupes d'individus, que ce n'est pas du côté de la médecine qu'il faut regarder, mais de l'agronomie et de la biologie (par exemple, Ronald Fischer) où il était facile de comparer deux champs de maïs recevant deux engrais différents. Ce n'est pas par hasard que les meilleurs livres de statistique dont a pu s'inspirer Van der Maren au moment de sa thèse (soutenue en 1971) étaient écrits par Pierre Dagnelie, professeur à la Faculté des sciences agronomiques de Gembloux (Belgique). Ce ne sera qu'autour des années 1925-1930, et pour des raisons de reconnaissance d'un plus haut statut universitaire, que des chercheurs en pédagogie, comme Raymond Buyse (1935), compareront leur pratique à celle de la médecine et se réclameront d'une parenté avec Claude Bernard et son « Introduction à l'étude de la médecine expérimentale » (1865). Notons qu'alors il s'agissait surtout de comparaisons systématiquement menées entre des observations de traitements, comme l'avait fait Ignaz Semmelweiss dans son étude de la fièvre puerpérale (1861), en comparant la mortalité des parturientes accouchées par un médecin qui s'était lavé les mains en sortant de la salle d'autopsie et de celles accouchées par d'autres médecins qui ne se les étaient pas lavées.

2. Preuve scientifique et preuve statistique

Une autre confusion fréquemment lue porte sur la différence entre preuve scientifique et preuve statistique. Un article d'André Tricot (2017), comparant la recherche scientifique productrice de connaissances et la recherche ergonomique fournissant des solutions, est très clair à ce sujet. Voici ce qu'il nous dit à propos des connaissances produites par la recherche scientifique:

Une connaissance [scientifique] est relativement autonome du lieu et du temps où elle a été créée. Elle devient scientifique quand elle est en outre relativement indépendante de la personne qui l'a produite et que cette personne a été relativement claire sur les conditions de production et de validité de cette connaissance : dans les mêmes conditions et avec une méthode identique, une autre personne aurait censément produit la même connaissance et en aurait identifié les mêmes limites. L'initiative de la production de connaissances scientifiques revient aux chercheurs, qui, en fonction de l'état des connaissances de leur domaine, décident que telle ou telle question constitue un véritable enjeu de connaissance. Le fait qu'une connaissance scientifique devienne un jour utile « en pratique » est difficile à anticiper.

Rien dans le texte de Tricot ne fait référence aux statistiques. La preuve scientifique exige seulement 1°, que l'utilisation de la méthodologie soit indépendante du chercheur et des sujets produisant les données (d'où l'exigence du double aveugle, Van der Maren, 1977), 2° que le résultat attendu comme point de bascule entre la confirmation et l'infirmité soit déterminé avant la collecte des données et 3° que la méthodologie soit complètement et clairement documentée de telle sorte que la recherche puisse être reproduite et aboutisse à des résultats semblables. Exprimée autrement, une connaissance scientifique ne peut être construite sur le rapport d'une observation perçue une seule fois par une personne intéressée aux effets de cette observation, que cette observation soit techniquement construite ou non.

La preuve statistique de la confirmation ou de l'infirmité d'une hypothèse, pouvant éventuellement constituer une connaissance, implique avant tout que l'utilisation des statistiques suive les exigences de la méthode scientifique. Comme il s'agit de répéter la méthode de collecte des observations sur plusieurs événements (ou sujets) dans des conditions bien identifiées, les statistiques « *descriptives* » sont essentiellement des résumés (une moyenne et un écart type, une médiane et un percentile) d'ensembles d'observations et de comparaisons des résumés. Notons que, pour chacun des résumés qui indique la centralité d'un ensemble de données (moyenne, médiane) un autre résumé fournit une indication de la dispersion ou des variations autour de cette centralité (écart type, quartile). Autrement dit, il est faux de prétendre, comme nous l'avons lu, que ces statistiques ne tiennent pas compte de la diversité des sujets et de leurs comportements.

Par ailleurs, la statistique « *inférentielle* » est un outil probabiliste, une sorte de pari : si tel résultat (par exemple, une différence entre une intervention et une autre) est obtenu avec une probabilité minimale admise par la communauté des chercheurs, on peut faire le pari qu'elle n'est probablement pas attribuable à une erreur d'échantillonnage. La probabilité

associée à un résultat statistique n'indique pas que ce résultat est significatif au sens où il constituerait la preuve pertinente et vraie de quelque hypothèse. La probabilité associée à un résultat indique seulement celle d'obtenir le résultat observé à condition que les échantillons aient été bâtis de manière aléatoire ou que l'on ait la quasi-certitude qu'ils sont équivalents pour toutes les dimensions pertinentes, c'est-à-dire pour les variables en jeu dans l'hypothèse testée. La probabilité observée, le petit « $p < .05$ », donne seulement la signification statistique (probabiliste) et non pas la signification pratique : il ne donne pas la validité pratique du résultat. C'est la taille du résultat observé, par exemple l'ampleur de la différence des performances entre deux groupes à une épreuve, qui peut être interprétée comme indiquant ou non un effet. Mais là aussi, si on désire faire œuvre scientifique, la taille de l'écart qui permet de trancher doit être déterminée à l'avance. On ne peut pas décider après coup : le chercheur ne peut pas, comme le politicien qui se trouve devant un résultat moindre qu'attendu, s'en tirer en disant que c'est quand même une tendance encourageante et positive.

3. Pourquoi la recherche expérimentale ne serait-elle pas possible en éducation ?

Une autre désinformation largement véhiculée consiste à prétendre que, pour pouvoir effectuer une recherche expérimentale en éducation, il faudrait que tous les enseignants impliqués dans une recherche fassent exactement le même enseignement en exécutant de manière identique le même procédé ou dispositif d'enseignement. Or, les plans expérimentaux ou quasi expérimentaux ne postulent pas l'homogénéité des interventions. Ce sont seulement certains modèles d'analyse statistique paramétrique qui réclament l'homogénéité des variances. Or, les statistiques non paramétriques n'exigent pas l'homogénéité des variances, pas plus que des échelles d'intervalles ou de rapport. Il s'agit, bien sûr, de ne pas comparer des oranges avec des pommes, mais rien n'empêche, pour garder cet exemple, que l'on puisse comparer différentes sortes de pommes pour qu'une compote soit délicieuse. De plus, certains plans d'expérience, comme « les plans avec groupes appariés et anticipation » et la « comparaison des performances a posteriori pour un plan à six groupes avec anticipation » (Van der Maren, 1996, pp 228-232) n'ont pas cette exigence a priori. On y compare, pour chacun des groupes expérimentaux (avec aléatorisation⁴) ou quasi expérimentaux (sans aléatorisation), les performances des individus (ou sujets) qui témoignent de caractéristiques (habiletés, compétences) semblables avant l'intervention. On constitue ainsi des paires, des trios, des quatuors de sujet semblables alors qu'ils sont répartis dans les différents groupes. Ce sont les sujets semblables issus de chaque groupe, ou de chaque situation expérimentale, qui sont comparés entre eux et non pas les groupes. La question alors est: est-ce que la plupart (nombre à déterminer à l'avance) des sujets du groupe expérimental ont un gain (ampleur à déterminer à l'avance) supérieur à la plupart des sujets du groupe contrôle?

4. Les données probantes sont-elles un dictat ?

Notons aussi que dans les situations complexes ou avec les individus complexes (composés d'éléments en interaction, mais de nature différente– incommensurables), comme le sont les êtres humains, aucune théorie éprouvée par la recherche expérimentale ou aucune

⁴ Randomisation est l'anglicisme correspondant au mot français aléatorisation, soit : répartition au hasard.

pratique conforme aux données probantes n'est absolue : cela reste la meilleure hypothèse, celle qui résiste le mieux aux contradictions. Et lorsqu'il y a des faits qui infirment l'hypothèse ou le dispositif attesté, on parle de la réduction ou de la délimitation de son champ d'application par et pour la classe des sujets ou des événements qui n'obéissent pas à la prescription générale attestée par les données expérimentales ou par les données probantes.

Prenons un exemple vécu par l'un d'entre nous. Dans la période actuelle où les ordres médicaux craignent une surprescription d'opioïdes, le modèle de traitement préconisé consiste à donner aux victimes de sciatique très douloureuse de l'acétaminophène (ou paracétamol en Europe; marque commerciale : Tylenol) et de la prégabaline (marque commerciale : Lyrica). C'est la prescription que suivit le médecin des urgences. Or, en plus des effets secondaires désagréables, mais tolérables, dont la somnolence et des troubles digestifs, certains individus réagissent très mal au Lyrica qui, chez eux, peut provoquer des agitations incontrôlables (mouvements des bras et des jambes) qui peuvent durer et fortement perturber le sommeil. Le médecin de famille a très vite corrigé le traitement en revenant, de manière très contrôlée, à un opioïde léger. Notons bien que la revue québécoise « Psychomédia », et la revue française « Prescrire » avaient attiré l'attention des médecins sur les effets négatifs du Lyrica et le peu de gain par rapport à un placebo dès le 1 juillet 2014 et l'avaient répété le 26 mars 2017. Il y a donc une littérature disponible sur ce traitement qui questionne une prescription influencée par une politique (lutte contre les opioïdes), ce qui montre bien la place et le rôle indispensable d'un Institut National⁵ d'Excellence en Santé et en Services Sociaux, ici, sur la santé. C'est aussi ce qu'on attend en éducation de l'Institut National d'Excellence en Éducation⁶.

5. La nécessité des données probantes bannit-elle toute autre forme de recherche ?

Là où certaines déclarations pourraient être malhonnêtes (espérons que non) ou indiqueraient une lecture un peu trop rapide et sans compréhension de l'écrit qu'ils critiquent, c'est quand, comme dans un titre, des auteurs prétendent que le texte de Gauthier, Bissonnette et Van der Maren (2023) met « à l'index un pan tout entier de la recherche en éducation au Québec ». Or, ce que ce texte qu'ils critiquent dit, c'est qu'après avoir examiné différents rapports de signification entre les mots théorie et pratique, notre texte se concentre, ensuite et seulement, sur deux des rapports (Vrai et Efficacité) car les autres rapports, dont l'Utilité qui renvoie à l'application des résultats des recherches fondamentales, ne sont pas concernées par ce que l'on entend, dans ce texte, par des « données probantes ».

⁵ Un collègue nous a dit récemment : "... dommage que tous les enseignants universitaires en éducation [quelle que soit leur discipline] n'aient pas bénéficié d'une formation inspirée du « Référentiel de formation des chercheuses et chercheurs francophones en éducation » (2019) ; au moins ils sauraient lire et comprendre, de manière critique et nuancée, l'intérêt, la signification et la pertinence spécifique des différentes modalités de la recherche en éducation, qu'elles soient quantitatives ou qualitatives".

⁶ Espérons que les gouvernements, l'actuel et ceux qui suivront, n'en feront pas une auberge espagnole et lui maintiendront ses objectifs initiaux que nous avons soutenus.

Cela ne veut pas dire que les recherches appliquées, dont font partie les recherches évaluatives à côté des recherches-actions et autres recherches participatives, y compris les analyses phénoménologiques, sont inutiles et n'ont pas de valeur. Elles ont toutes une valeur, mais leurs validités internes, et surtout externes, sont différentes de celles de la recherche évaluative comparant plusieurs interventions. Ce que le texte qu'ils critiquent dit c'est que la recherche évaluative a une fonction importante, sinon essentielle, *à un moment donné*, c'est-à-dire après que les responsables de l'enseignement et le public se soient entendus sur les finalités de l'institution éducative, et pour *une tâche précise* : la gestion des méthodes et des dispositifs d'enseignement à plus ou moins grande échelle (une école, un centre de Service scolaire, une région, une province, un état) en vue d'atteindre les finalités préalablement choisies. Puisqu'il s'agit d'une application à grande échelle, la démarche du choix se base sur le modèle de la recherche de connaissances scientifiques et non pas, comme l'ergonomie, sur la recherche de solutions pratiques pour des situations précises et limitées. Lorsqu'il s'agit de fonder un choix de dispositifs en vue de la réalisation d'une politique ou d'une philosophie pédagogique, c'est-à-dire d'une finalité, la première étape est la délimitation de cette finalité. Une fois qu'un accord, un consensus ou un choix politique a été fait, une saine gestion doit faire l'hypothèse du meilleur dispositif pour l'atteindre; et c'est là que la recherche évaluative avec des groupes quasi expérimentaux devient indispensable. Comme les experts de l'éducation ont la plupart du temps des intuitions divergentes, des présupposés idéologiques différents et des propositions pédagogiques originales et généreuses auxquels ils adhèrent et promeuvent, l'imposition d'une méthodologie d'enseignement à toute une organisation pédagogique nécessite une évaluation au moins quasi expérimentale.

Cela ne veut pas dire que les modalités non expérimentales de recherche ne peuvent pas jouer un rôle important, mais en aval et en amont d'une recherche évaluative aboutissant à des données probantes. Sigmund Freud a dit qu'il y avait 3 métiers impossibles : gouverner, enseigner et soigner. Ce sont trois métiers où le résultat obtenu est toujours inférieur au résultat attendu et, par conséquent, il est toujours une source d'insatisfaction. Il est clair que le déni est un des mécanismes de défense pour les professionnels de ces métiers, entre autres pour les promoteurs d'une mode pédagogique ou d'un dispositif innovant. Nous avons rarement entendu un politicien, en fin de mandat gouvernemental, dire qu'il n'avait pas réussi à améliorer la situation de la nation, bien plus que ses prédécesseurs (ou celui de la province voisine). Nous n'avons jamais vu un patient qui ne mourrait pas un jour, la vie étant une maladie mortelle et la mort un manque de savoir-vivre, quoique puissent faire le médecin et l'infirmière. Nous ne connaissons pas une classe où tous les élèves obtiennent une note parfaite; nous sommes toujours déçus de ne pas avoir pu conduire un élève à une meilleure réussite.

Mais allez donc dire à certains professeurs d'université en éducation que le projet généreux et les dispositifs géniaux auxquels leur intuition et leur réflexion les ont conduits ne marchent pas mieux, ne sont pas aussi efficaces que ce que d'autres dispositifs font !

À moins qu'une maladie ne réduise son énergie et sa lucidité, ou qu'il soit narcissique et trouve des causes externes à la permanence des difficultés qu'il rencontre dans sa classe, l'enseignant d'expérience réagira sur le coup, sinon après sa journée ou sa semaine. Il

reviendra sur l'échec de son intervention et réfléchira (seul ou, mieux, avec quelques collègues) aux améliorations qu'il pourrait apporter pour en réduire la difficulté. En général, cette rétroaction auto-formative ne touchera que quelques élèves et cela momentanément. C'est là qu'un parallélisme peut être osé avec ce que fait l'ergonomie : trouver une solution locale à ce qui ne marche pas, parfois même malgré les recherches, malgré les données probantes, soit ce qui a été prouvé comme efficace dans la majorité des cas, mais qui ne fonctionne pas avec ce cas particulier.

Malheureusement, certains enseignants universitaires peuvent persister à enseigner une des intuitions et des innovations généreuses, mais sans efficacité pratique, à une volée de futurs enseignants d'école, parce qu'ils ne voient pas (ou ne veulent pas voir) les conséquences de leur enseignement autant que la professeure d'école confrontée immédiatement à la non-réussite, à la résistance cognitive de certains de ses élèves. Quand l'évaluation du dispositif imaginé avec génie et générosité par certains leur montre les limites de l'efficacité espérée, s'ils sont lucides et vivent une attitude scientifique* (doute, scepticisme, ouverture aux débats, à l'écoute, à la polémique - Van der Maren, 2014, p.17), ils vont retourner à l'examen critique des théories ou de l'intuition à l'origine de l'innovation proposée, réviser le dispositif, entreprendre une nouvelle recherche de développement ou une enquête participative pour mieux identifier, par exemple, les conditions matérielles, organisationnelles ou psychologiques qui sont un obstacle au succès. Là, apparaît la pertinence de mener une enquête pour identifier la culture des écoles visées, la culture du milieu social et familial des élèves, les possibilités de collaboration entre la direction et les enseignants et des enseignants entre eux, ainsi que les ressources matérielles disponibles et le temps de formation, en plus de la tâche habituelle, que le projet exigerait. C'est là que des recherches tout à fait qualitatives, quasi ethnographiques, ont leur intérêt pour identifier l'ensemble des éléments du système classe qui résistent aux interventions de l'enseignant. C'est là aussi que des enquêtes phénoménologiques, comme les entrevues d'explicitation, sont utiles pour comprendre en quoi les ressentiments, les injonctions introjectées, les règlements déviés ou contournés, les fictions dirigeantes et les héritages culturels ont un impact sur les fonctionnements du métier de l'enseignant.

6. Preuve culturelle de l'efficience, pas de l'efficacité.

Certains auteurs ont ici évoqué ce que Sensevy (2022) appelle la preuve culturelle, qui n'est pas une preuve de l'efficacité d'un dispositif mais seulement, et tout au plus peut-être, un ensemble d'indicateurs de son efficience au sens où Le Moigne (1977) a distingué l'efficience de l'efficacité. Pour faire court, l'efficacité est une mesure (souvent quantitative) du coût par rapport aux bénéfices et pose la question du choix de la meilleure méthode pour atteindre plus rapidement et avec un moindre effort l'objectif (la finalité) que l'on s'est fixé. De son côté, l'efficience est une évaluation ou bien du respect de normes professionnelles, d'une éthique, d'un cahier des charges ou bien, comme l'écrit Sensevy, du respect du savoir du connaisseur, ce qui équivaut au savoir d'un maître (comme dans la formation par compagnonnage), d'un spécialiste du domaine qui accompagne, guide et évalue l'apprenant. En bref, l'efficience est une mesure du bien faire les choses. Encore faut-il s'entendre sur ce qu'est « bien faire les choses » et cela dépend de la désignation de ceux

qui décident de ce qui est bien fait : une habitude commune, la culture d'une nation, une tradition éducative, la culture d'une école de pensée pédagogique? On y reviendra plus loin.

Dans certains cas, l'efficience et l'efficacité peuvent être en contradiction. C'est ce que l'on voit, par exemple, dans la construction ou la réfection des routes où le contrat est donné au plus bas soumissionnaire (efficacité ?) au détriment de la qualité du revêtement (efficience) qui exigerait des surplus, des dépassements de coût, par rapport au plus bas soumissionnaire, pour que la surface de roulement soit solide assez longtemps et sans défauts. Ce qui correspond aussi, assez souvent, à l'écart que l'on peut observer entre les perspectives technologiques, qui visent à ce que cela fonctionne, et les perspectives artisanales qui visent d'abord à ce que cela soit bien fait et même, dans certains cas, à ce que le résultat soit beau quelle que soit la manière et le temps ou l'argent exigés pour y arriver. Par exemple, si des études montrent l'effet positif de la diminution du nombre d'élèves par classe, elles montrent aussi que le coût est élevé⁷.

Le tableau⁸ ci-dessous montre les divergences entre différentes perspectives sous le titre des questions aux paradigmes. Nous utilisons ici la signification du mot paradigme tel que présentée par De Bruyne et coll. (1971) qui le définissent comme un ensemble d'éléments partagés par une communauté dans différents domaines qui recourent à des connaissances à des savoirs, à des habiletés: domaine dogmatique (foi et morale religieuses), épistémique (recherche de connaissances scientifiques), polytechniques (production et gestion), artisanale (fabrication et intervention) et artistique (création). Cet ensemble partagé par une profession, un métier, une corporation compte 5 éléments : 1°, une définition de ce qu'est leur objet, soit leur ontologie; 2°, les règles ou principes doivent respecter leurs activités, soit leur sémantique; 3°, les modalités instrumentales qui insèrent leurs activités dans la réalité, soit leurs techniques; 4°, les formes de communication qui assurent leur diffusion, soit leur rhétorique; 5°, la caractéristique visée par leurs activités qui permet de les évaluer, soit leurs exigences externes.

Lorsqu'on considère les 3 professions au centre de ce tableau et qu'on les situe dans le champ de l'éducation, soit la recherche de connaissances (épistémique), la gestion des systèmes et institutions (technique) et l'action d'instruire, de former des élèves (artisanat), on voit bien que les exigences externes, tout comme les autres dimensions de leur paradigme diffèrent : on passe de la vraisemblance ou de la signifiante pour les recherches scientifiques ou herméneutiques, à l'efficacité pour la gestion, à l'efficience pour l'instruction. Pourquoi cette différence d'exigence pour la gestion des institutions et l'instruction des élèves ? Parce que dans leurs classes, les enseignants se retrouvent en face d'une portion importante d'élèves avec lesquels ils doivent adapter, modifier, ajuster les recommandations, les programmes, les méthodes et didactiques recommandées par les gestionnaires. Autrement dit, dans leurs classes, les enseignants sont contraints de bien

⁷ Malheureusement, c'est surtout au coût que les gouvernements sont sensibles quand il s'agit de l'éducation ou de la santé. Par ailleurs, l'efficacité en termes de coût - bénéfice est essentiel pour les pays moins riches que le Canada; voir, par exemple, rapport de la Banque Mondiale : Cost effective approaches to improve global learning.

⁸ Le tableau des questions paradigmatiques aux disciplines est le tableau complet présenté lors d'une conférence, du tableau 1.1, page 30, de Van der Maren, J.M., (2014), La recherche appliquée pour les professionnels, De Boeck, Bruxelles.

faire ce qu'ils ou elles peuvent faire avec les élèves qu'ils ont dans le cadre des programmes et dispositifs efficaces pour la majorité de leurs élèves, selon les prescriptions des gestionnaires. Comme dans l'exemple du médecin, cité plus haut, qui déroge de la prescription du Lyrica pour un patient sciatique qui y est allergique. Il a moins été efficace du point de vue du système de santé, mais il a été plus efficace du point de vue du patient. C'est l'exception à laquelle les praticiens de la santé comme ceux de l'éducation sont contraints tout en respectant le cadre général des interventions promues par les recherches évaluatives sur les procédures de soins et sur les processus d'enseignement, quand les gestionnaires s'appuient sur la science plutôt que sur une culture locale. Notons, en passant, que Patricia Pitcher, qui fit sa thèse de doctorat sous la direction d'Henry Mintzberg à l'Université McGill avant de devenir professeure aux HEC, classe les enseignants non pas comme des artistes ou des technocrates, mais comme des artisans (1997). Comme tous les enseignants, en formation des maîtres nous sommes aussi des artisans, alors que les chercheurs se pensent parfois, sinon souvent des artistes (des innovateurs).

QUESTIONS PARADIGMATIQUES					
Pôles	DOGMATIQUES RELIGIONS	ÉPISTÉMIQUES NOMOTHÉTIQUES (causalité) KATANOÉTIQUES (sens)	TECHNIQUES TECHNOLOGIQUES (production, intervention)	FORMATIVES - ARTISANALES (action, fabrication)	ARTISTIQUES
<i>Ontologique</i>	FOI : quelle contenu de croyance incontestable et instable offrir en contrepartie des exigences morales ?	Épistémologie : Quelle est la réponse admissible aux questions du rapport des hypothèses à la réalité ?	Fonctionnalité : le procédé, le produit fonctionne-t-il de manière efficace (moindre coût, moindre erreur) ?	Éthique : De valeurs, de quelles responsabilités témoigner qui organisent et rendent légitime l'action ?	Artistique : quelles variations dérangeantes et enchanteuses produire pour être (et rester) unique et admiré ?
<i>Sémantique</i>	Exégétique : quel forme d'hermétiques sacralisantes et d'interprétations doit soutenir la prédication	Théorique : Quelle précision et cohérence et consistance doivent avoir entre eux les concepts pour que les discours soient admis comme théories ?	Opérativité : les instructions sont-elles claires, simples, faciles, précises et complètes pour assurer le succès de l'utilisation ?	Esthétique : Quelles manières d'agir, quelles règles de l'art respecter pour que l'action soit reconnue comme professionnelle ?	Esthétique : quelle organisation, arrangement, dynamique icono-scénique exprime la limite personnelle des normes en suscitant ou en maintenant l'audience ?
<i>Technique</i>	Ritualisation : quelles célébrations, quels ritualisations, soutiennent l'adhésion et la participation.	Instrumentation : Quelles inscriptions (données, analyse, traitements) sont efficaces, permettent de fournir et de soutenir des hypothèses ?	Cahier des charges et Accessibilité : les ressources et les moyens indispensables à la mise en oeuvre sont-ils accessibles à un coût acceptable ?	Pragmatique : Quels usages spécifiques des outils généraux et spécifiques du métier permettent l'efficacité et l'efficience de l'action ?	Quelles connaissances des matériaux et quelles compétences techniques savoir maîtriser pour les exploiter dans des variations personnelles étonnantes et attirantes ?
<i>Rhétorique</i>	Quels contenu, formes, et structures liturgiques permet de soutenir la foi et la soumission morale ?	Communication : Quelles formes de présentation doivent exprimer et soutenir les hypothèses (formules, tableaux, graphiques, illustrations par des extraits, par des vidéos) ?	Publicité : Les arguments de la présentation rendent-ils le produit désirable et acceptable, adoptable ?	Scénarisation : Quelles manières d'intervenir, de montrer l'action et d'en parler indiquent sa recevabilité et sa contribution à la respectabilité du métier ?	Exhibition : quelles lieux, moments et organisations des présentations mettre en place pour exhiber une production à la fois personnelle et culturelle (reconnue comme telle) ?
<i>Exigence</i>	croyance - assurance - exigence	Vraisemblance / Signifiante	Efficacité	Efficience	Étonnement enchanteur

7. Évidentialité, preuve culturelle, preuve basée sur la pratique

Cela nous ramène à l'article de Sensevy, utilisé comme argument par Pulido et Allaire, dans lequel, sans nous laisser aller à une analyse critique que mériterait ce texte un peu confus, nous pointons seulement ce qu'il veut dire en parlant de la preuve culturelle et de ses limites. Cet auteur recourt d'abord à de nombreux dictionnaires qui examinent le sens courant des mots réalité, preuve, évidence, et emprunte à la linguistique le mot

« évidentialité », sans que l'on voie comment les dictionnaires ou vocabulaires spécialisés, en particulier en éducation et chez les méthodologues, définissent dans nos disciplines, ces mots. Selon Cécile Barbet et Louis de Saussure (2012), des linguistes, l'évidentialité concerne, dans un sens restreint, « l'indication par le locuteur de la source de l'information véhiculée par son énoncé », alors que dans un sens plus large elle concerne l'indication de la fiabilité de l'information communiquée. On est loin de la preuve. L'exemple emprunté par Sensevy n'est pas sans rappeler, au Canada et au Québec, les langues algonquiennes qui, par l'ajout de préfixes et de suffixes, indiquent de manière plus précise des actions, des états ou des acteurs différents. La précision d'une information, de sa source ou de sa fiabilité ne sont pas preuve de l'existence d'un fait. Nous pouvons raconter que nous, personnes habituellement crédibles et âgées (donc expérimentées) avons vu un renard, alors que ce n'était qu'un chien. En avons-nous fait une photo assez nette (et non truquée par l'IA) pour en être certains ? Avons-nous vraiment pris la photo ou l'avons-nous piquée sur Internet ? Ce n'est pas plus parce qu'un pape (une personne respectable) a proclamé le dogme de l'Immaculée Conception qu'elle est un fait auquel nous adhérons même lorsqu'on nous dit que c'est le pape qui l'a dit. Ce n'est pas non plus parce que tous les catholiques l'admettent que c'est un fait vérifiable et prouvé.

Par ailleurs, il n'y a pas de preuve statistique : il y a des conventions statistiques qui soutiennent une preuve scientifique si les données traitées par des outils statistiques ont été recueillies en respectant les exigences de la recherche scientifique, telles que nous les avons listées plus haut. On ne peut donc pas opposer la preuve statistique à une « preuve basée sur la pratique » (practice-based evidence). Une pratique n'apporte jamais une preuve attestant de la validité d'une hypothèse. Mais elle peut apporter une contre-preuve, c'est-à-dire constituer un cas qui met en cause la généralisabilité d'une hypothèse devenue théorie. Encore faudrait-il que plusieurs cas pratiques, sans être l'effet d'un complot, apportent une même contre-preuve. Comme nous l'avons souligné, la contre-preuve répétée conduit la théorie à ne plus être généralisable, autrement dit, les cas contradictoires exigent que le domaine ou champ d'application de la théorie soit réduit en tenant compte des cas où elle ne s'applique pas.

Qu'est donc une « preuve culturelle » ? Pour Sensevy, c'est une preuve basée sur la pratique, basée sur le savoir d'expérience du praticien. Il ajoute que les preuves culturelles sont liées « à la vie pratique, produisant elle-même des évidences, de la certitude, sans lesquelles elle serait à la fois impossible et incommunicable ». Il semble oublier que ces évidences ne sont pas des preuves, mais des conventions transmises de génération en génération, comme des axiomes que l'on admet et que l'on partage sans pouvoir ni sans avoir à les prouver parce qu'ils sont nécessaires ou coutumiers. Il distingue en outre deux sortes de preuves culturelles : les preuves communes et les preuves des connaisseurs de pratique. Les preuves (dit-il au lieu de savoirs) communes, sont les faits, les événements et les pratiques éducatives qui sont des exemples évidents qui font l'unanimité de ceux et celles qui ont une expérience suffisante de la pratique. Les preuves des connaisseurs de pratique apparaissent comme des « systèmes de connaissances (savoirs) avérées et reconnus par le collectif qui reconnaît et transmet les arts de faire ». Ni l'une ni l'autre ne sont des preuves d'efficacité : elles sont des indices de conformité à la pensée commune, à une culture ou à une école de pensée pédagogique ou à une institution de formation professionnelle. Et la plupart du

temps, elles sont des indicateurs d'efficacité aux yeux du commun d'une population ou d'un ordre professionnel. La plupart des parents ont un savoir culturel commun, ayant passé au moins 11 années à l'école, et cette expérience, plus ou moins commune, leur permet, pensent-ils, de juger d'un enseignement bien fait. Mais la conformité à leur savoir d'expérience issu de leur passé comme élève ne peut pas être une preuve du bien ou du mal faire effectif de l'enseignement actuel à partir de ce que leur rapporte leurs enfants de l'enseignement reçu. Quant à la preuve culturelle du connaisseur de la pratique, elle n'est que l'observation des indices plus ou moins explicites de la conformité d'une pratique à celle recommandée par un enseignement d'une école de pensée pédagogique ou par une formation donnée dans une institution ou un ordre professionnel. En aucun cas, il ne s'agit de preuves scientifiques d'efficacité pratique, elles ne sont essentiellement que les indices sur lesquels peuvent se baser des parents, des gestionnaires, de politiciens pour estimer la conformité à leurs attentes. Mais d'où viennent leurs attentes ? Malheureusement, pas toujours de l'évaluation scientifique de l'efficacité ni même de la définition rigoureuse de ce que pourrait être, pour une population définie, l'efficacité.

Concluons cette analyse des prétentions de l'article de Sensevy en soulignant que son désir de condamner les pseudo preuves statistiques par des preuves culturelles est une argumentation fallacieuse dans la mesure où il n'y a de preuve statistique que là où les règles de la recherche scientifique ont été respectées et qu'il n'y a donc pas de preuve statistique en soi. Il n'y a pas plus de preuves culturelles, ce qu'il entend par là n'étant que des indices, souvent peu discriminants, mais partagés par certains groupes culturels ou professionnels, pour juger de la conformité des productions à leurs attentes fondées sur des habitudes, sur des traditions, sur des idéologies.

8. Le rôle des autres formes de recherche en éducation

Revenons à la recherche évaluative et aux données probantes qui n'excluent pas du tout les autres formes de recherche appliquée. Ces autres formes de recherche se trouvent avant et après celles qui aboutissent à l'efficacité d'une méthode d'enseignement, d'un dispositif. Avant ce sont celles qui, une fois les finalités de l'éducation définies, s'interrogent sur les possibilités de réalisation d'une intuition, sur ses conditions de succès et les recherches qui tentent encore d'améliorer les méthodes et les dispositifs après qu'une évaluation scientifiquement conduite en ait montré les possibilités et les limites.

Par ailleurs, nous pensons aussi que la recherche évaluative suppose une comparaison entre l'avant et l'après d'une intervention, avant et après la mise en action d'un dispositif. C'est la comparaison temporelle simple. Mais le contexte pourrait interférer favorablement ou défavorablement. C'est pourquoi on va privilégier plusieurs observations ou mesures avant, pendant et après l'intervention ou l'application du dispositif. Cela permettra de voir si le changement (un apprentissage) observé n'était déjà pas présent avant l'intervention, si l'intervention l'a accéléré, stabilisé ou ralenti et s'il s'est maintenu après l'intervention, au même rythme qu'avant ou avec un rythme accéléré par l'intervention ou si l'intervention a eu peu d'effet à plus ou moins long terme. Ces exigences méthodologiques expliquent que ces recherches évaluatives ne sont si simples à organiser ni peu coûteuses en temps, en personnel et donc en financement. Ces exigences expliquent donc pourquoi elles ne

foisonnent pas et pourquoi ce sont les études avant l'évaluation de l'efficacité (par exemple, les recherches d'innovation, les recherches-actions) ou après l'évaluation (les recherches de développement à partir de recherche d'innovation pour améliorer des implantations locales) qui sont les plus répandues. Ces recherches, avant ou après la recherche évaluative expérimentale, produisent des observables plus rapidement et ont moins d'exigences en personnels et financières. Par ailleurs, si ces observables plus rapides semblent signaler une certaine efficience en rencontrant certaines attentes, non seulement les gestionnaires pourraient s'en contenter avant d'obtenir les résultats d'une évaluation expérimentale, mais en plus elles peuvent fournir des indications sur les conditions nécessaires à la mise en place d'une évaluation expérimentale d'envergure.

9. Les études de cas et les études pilotes.

Les études de cas et les études pilotes ne permettent pas la généralisation tout en étant nécessaires lorsque les connaissances scientifiques ne sont pas suffisamment disponibles. La comparaison ante-post dans une classe par un enseignant, valide et pleine de renseignements pour lui, est insuffisante pour être généralisée à d'autres. On dira qu'elle a une bonne validité interne, mais une faible validité externe. Bien évidemment, les collègues de cet enseignant pourraient s'en inspirer, mais c'est trop localisé pour qu'une direction puisse fortement la recommander aux autres. La recherche évaluative qui vise tant une bonne validité interne qu'une validité externe, la généralisation d'une recommandation, exige que soient comparés entre eux les gains relatifs de différentes interventions. Si plusieurs applications d'une même méthode ou d'un dispositif identique dans des classes différentes, ce qui imposera de facto des ajustements locaux, montrent des gains qui vont dans le même sens, malgré les ajustements locaux, alors la validité externe augmente et une recommandation de la méthode ou du dispositif le plus efficace devient mieux fondée, et justifiée. Ainsi, nous pensons que les résultats de plusieurs recherches comparables, mais pas identiques, dans des lieux et/ou des époques successives ou éloignées, ont plus de validité externe et donc plus de généralisabilité qu'une seule recherche faite dans une seule situation, surtout si elle est unique comme les sont généralement les recherches menées dans les écoles et les classes « volontaires », ce qui comporte—de nombreux biais (Rosenthal, 1969; Rosenthal et Jacobson, 1971). Aussi, on ne peut pas se contenter ni d'une étude de cas, ni d'une étude pilote menée dans plusieurs classes et même dans plusieurs écoles parce que les études pilotes sont trop proches des promoteurs et qu'en général les études pilotes bénéficient de conditions favorables (publicité, volontariat, formation spéciale et financement, reconnaissance d'un statut à ceux qui y participent) que n'auront pas les enseignants et les classes lors de la généralisation à un plus grand ensemble. L'étude de cas et les études pilotes ont cependant tout leur sens lorsqu'on les utilise pour identifier les conditions nécessaires à la réussite d'une implantation généralisable et à la vérification, par une recherche évaluative expérimentale comportant les mêmes conditions que les institutions (Écoles, Centre de service, Ministères) pourront effectivement leur accorder.

Cela étant, il est normal que les recherches appliquées, les recherches qui visent l'utilité tant que cette dernière n'est pas démontrée, foisonnent parce que l'échec relatif, mais inéluctable, de la situation d'enseignement persiste. De plus, la mode contemporaine et consumériste de l'innovation pour « l'innovant » pousse le CV des universitaires, quitte à

occulter ce que les prédécesseurs ont fait de plus ou moins bien réussi. Cependant, comme on l'a indiqué précédemment, étant donné les difficultés techniques de réaliser des recherches évaluatives expérimentales d'envergure régionale ou nationale et les résistances auxquelles elles font face, il est aussi normal qu'elles soient plus rares. D'autant plus qu'elles impliquent plus de personnel, sur une longue période et coûtent plus cher. Or, comme les politiciens, lorsqu'ils ne sont pas « des Hommes d'État », gèrent le moment, l'économie immédiate et leur réélection, les plans de recherche évaluative septennaux et coûteux n'ont pas leur faveur, sont peu encouragés et subventionnés.

10. L'Éducation, ambiguïté entre idéologie et science, entre continuité et course à l'innovation.

Enfin, Evans, en 2007, a décrit la praxis comme étant « à la fois une action pensée et finalisée autant que finalité pensée et agie ». Ce qui nous renvoie à la situation à l'origine de la pédologie : la pédagogie scientifique était et a toujours été un compromis entre science et idéologie. Comme le texte évoqué (Gauthier & coll. 2023) le dit, ce n'est pas une théorie scientifique qui peut générer des finalités, des valeurs, des objectifs institutionnels. L'école a de tout temps intéressé les religions. Les gouvernements, les politiciens, y sont passés et s'y sont aussi fort intéressés une fois qu'ils ont pris le relais des religions pour les subventionner et faire de la fréquentation des écoles une obligation légale. De plus, tous ceux qui ont fréquenté les écoles ont souvent la science infuse de l'enseignement : ils connaissent l'école puisqu'ils y sont allés. En ce début du XXI^e siècle, le problème est qu'avec le déclin des religions catholique et protestante traditionnelles, qui homogénéisaient la culture des populations Ouest-Européennes et Nord-Américaines, un foisonnement d'idéologies a été favorisé par le multiculturalisme. Cela rebondit sur l'école, sur les conceptions que les pédagogues, les psychologues, les sociologues, les philosophes et les politiciens peuvent avoir du rôle de l'école, souvent devenue un fourre-tout, et sur les pédagogies, encore que chacun prétende souvent que la sienne est la meilleure pour les enfants.

Les facultés forment, non seulement des enseignants d'école, mais aussi des futurs professeurs d'université formateurs d'enseignants d'école. Ces facultés vivent la même multidisciplinarité, les mêmes diversités idéologiques et donc pédagogiques, du moins au niveau des discours qu'ils délivrent aux futurs enseignants des écoles. Et parce qu'ils sont professeurs d'université, ces universitaires des facultés et départements (de sciences ?) d'éducation se réclament en outre de la liberté académique, ce qu'en réalité ils n'ont pas. Tant qu'ils forment des futurs enseignants d'école, leur liberté n'est pas une liberté de parole, d'opinion. Ils sont, qu'ils le veuillent ou non, contraints à respecter un programme et les contenus du programme. Lorsque l'un de nous donnait un cours de didactique générale, il ne pouvait pas enseigner l'astrologie et la cartomancie. Quand il donnait des cours de méthodes de recherche en éducation, il ne pouvait ni donner des cours de cuisine ni donner des cours sur les sondages qui sont des objets d'études en sociologie et en sciences politiques, mais pas en éducation. Lorsque, comme professeur-chercheur dans une université nous faisons de la recherche fondamentale sur les processus d'enseignement et d'apprentissage, nous disposons d'une relative liberté académique. Mais, dans le cas de l'enseignement aux futurs enseignants des écoles, intervenant dans la formation à une

profession, nous sommes des hauts fonctionnaires de l'État qui nous paie via les subventions accordées à nos institutions et notre éthique professionnelle nous oblige à n'enseigner que ce qui est validé ou du moins ce qui fait consensus dans notre champ d'enseignement et dans le programme professionnel auquel nous participons. Quoique puissent en penser les syndicats lesquels, sur ces considérations, outrepassent souvent leurs fonctions légales. Notons qu'en recherche aussi, la liberté n'est que relative, car pour obtenir des subventions permettant la recherche, les candidats doivent, actuellement, souscrire à des exigences qui ne sont pas scientifiques, comme les exigences idéologiques de montrer qu'ils respectent les principes moraux EDI (Équité, Diversité, Inclusion) et les thématiques souhaitées par les gouvernements qui contrôlent les budgets des organismes subventionnaires. Ainsi, nous avons connu des collègues qui embarquaient sur des sujets à la mode pour lesquels ils n'avaient aucune compétence, mais qui étaient facilement financés. En outre, au Québec, les chercheurs doivent écrire leur demande de subvention en anglais s'ils veulent que la grande majorité des membres qui composent les comités de sélection et d'attribution des subventions fédérales comprennent leur projet et les évaluent correctement. Où es-tu, chère liberté académique ?

Terminons enfin par le paradoxe que nous avons vu lorsque certains universitaires, chercheurs et professeurs aux cycles supérieurs, mais aussi enseignants en formation des maîtres, réclament que tous s'entendent pour promouvoir les mêmes méthodes et dispositifs d'enseignement, alors qu'en même temps ils prétendent, comme chercheurs, devoir et pouvoir disposer de la liberté académique. Ce paradoxe tient sans doute de la situation de plus en plus ambiguë de la formation des enseignants dans les départements universitaires en éducation ou dans les Hautes Écoles Pédagogiques, alors que le statut envié ou imposé par les États se réalise surtout par le mythe de la « (grande) recherche » innovante et payante (par les subventions) plus que par l'enseignement et la formation efficace autant qu'efficace des maîtres des écoles. Cette ambiguïté n'existait pas dans les années 1960 à 2000 en Europe lorsque les Écoles Normales (ou Écoles Magistrales en Suisse) n'avaient qu'une fonction, la formation à l'enseignement des maîtres des écoles et que la maîtrise en éducation dans les universités formait les enseignants des Écoles Normales. Ce n'était aussi pas le cas de certaines universités au Québec tant qu'elles n'ont pas réclamé de former des chercheurs et chercheuses par les programmes de M.A et de Ph.D. Mais elles se sentirent sous-estimées, dévalorisées, alors qu'à notre sens il n'y a pas plus noble fonction dans une société que la formation des enseignants des Écoles. Pas de recherches savantes là, seulement des formations professionnelles inspirées des recherches faites ailleurs que dans les Écoles Normales, dispensatrices autant des connaissances scientifiques que des savoirs pédagogiques. Car la plupart des professeurs d'Écoles Normales avaient d'abord enseigné comme instituteurs (au primaire) ou professeurs (au secondaire) avant de faire la maîtrise pour accéder au poste d'enseignant des Écoles Normales; il fallut encore y ajouter le Ph.D. lorsque les formations des maîtres furent « universitarisées ». Les pratiques passées n'étaient pas si mauvaises que cela et l'histoire peut nous faire observer les dérives du présent.

Références

Altet, M. (2002). Une démarche de recherche sur la pratique enseignante ; l'analyse plurielle. *Revue française de pédagogie*, 138, p. 85093

Barbet, C. et de Saussure, L. (2012) Présentation : Modalité et évidentialité en français. *Langue Française*, numéro 173, pp 3 à 12.

Bernard, C. (1865), *Introduction à l'étude de la médecine expérimentale*. Réédition : Flammarion.

Bernard, C. (1867), *Principes de médecine expérimentale*. Émile Martinet Éditeur.

Buyse, R. (1935), *L'expérimentation en pédagogie*. Bruxelles : Lamertin.

Chrishman, O. (1896), *Paidologie : Entwurf zu einer Wissenschaft des Kindes*. Iena : Vopelins. (Pédologie, ébauche d'une science de l'enfant)

Dagnelie, P. (1998), *Statistique théorique et appliquée*. Bruxelles, DeBoeck.

Depaepe, Marc, (1999). La recherche expérimentale en éducation de 1890 à 1940: les processus historiques sous-jacents au-développement d'une discipline en Europe de l'Ouest et aux États-Unis. Dans Hostetter, R. et Schneuwly, B., *Le Paris des sciences de l'éducation*. DeBoeck Université, pp. 29-56 (voir ses références bibliographiques pour une étude historique plus détaillée).

Evans, R. Existing practice is not the template. *Educational Researcher*, 36, 553-559.

Gauthier, Clermont; Bissonnette, Steve et Van der Maren, Jean-Marie (2023). *La pertinence de l'utilisation des données probantes en éducation à la lumière des rapports théorie-pratique*. Manuscrit non publié, Université TÉLUQ, Département Éducation. Répertoire des publications en accès libre. Accessible sur : <https://r-libre.telug.ca/3064/f>

Johnson, C. (2019). *Land of Wolves*, Traduction française, Le pays des Loups, chez Gallmeister, 2023.

Le Moigne, J.-L. (1977). *Théorie du système général, théorie de la modélisation*. Paris : PUF.

Pitcher, P. (1997). *Artistes, artisans et technocrates dans nos organisations. Rêves, réalités et illusions du leadership*. Montréal, Éditions Québec-Amérique. (2e édition).

Pulido, L. Allaire S. (2023). Proposition pour définir et mobiliser les « pratiques efficaces reconnues par la recherche » In T. Laferrière, D. Savard, M-A. Éthier, Makdissi, H. (Eds.), *Le PL23 et l'INEE : excellence ou standardisation en éducation ? Réserves et propositions d'universitaires*. [En ligne]. Disponible : <https://drive.google.com/file/d/17LdIN78EQguRJA85im-GUWlfa1AGOBkA/view>.

Rosenthal, R. (1969). La participation volontaire. Dans Lemaine, G. et Lemaine J.M, *Psychologie sociale et expérimentation*. Paris : Mouton, pp. 71-79

Rosenthal, R. et Jacobson, L. (1971). *Pygmalion à l'École*. Tournai: Casterman.

Semmelweis, I.P. (1861). *Die Aetiologie, der Begriff und die Prophylaxis des Kindbettfiebers*. Pest-Wien-Leipzig, (Étiologie, concept et prophylaxie de la fièvre puerpérale).

Voir sur France culture un récit intéressant de sa vie et son œuvre.

<https://www.radiofrance.fr/franceculture/semmelweis-le-medecin-qui-tenta-d-imposer-le-lavage-de-mains-8212272>

Sensevy, G. (2022). Vers une épistémologie des preuves culturelles. *Éducation et didactique*, pp. 145-164. (En ligne)16-2, mis en ligne le 2 janvier 2024. URL: <http://journals.openedition.org/educationdidactique/10415> ; DOI : <http://doi.org/10.4000/educationdidactique.10415>

Schuyten, M.C. (1911), *La pédologie* (Synthèse). Gent : Vanderpoorten

Tricot, A. (2017). La connaissance et la solution. *Éducation et didactique*, 11-2, pp. 57-61. (En ligne) 11-2017 URL : <http://journals.openedition.org/educationdidactique/2725>.

Van der Maren, J.M. (1977). Le double aveugle contre Pygmalion : Éléments de psychosociologie de la recherche en éducation et méthodologie des plans. *Revue des sciences de l'éducation*, Vol.3 n.3, pp. 365-380.

Van der Maren, J.M. (1996), *Méthodes de recherche pour l'éducation*. Montréal : PUM et Bruxelles : De Boeck.

Van der Maren, J.M. (2014) *La recherche appliquée pour les professionnels*. Bruxelles : De Boeck.

Van der Maren, J.M., Brodeur, M., Gervais, F., Gilles, J.L., Voz, G., (2019), *Référentiel pour la formation des chercheuses et des chercheurs francophones en éducation*. Montréal : ADEREQ. Accès en ligne sur Papyrus, UdeM. Permalien : <https://doi.org/11866/21889>.

World Bank document (2020). *Cost-Effective Approaches to Improve Global Learning: What Does Recent Evidence Tell Us Are “Smart Buys” for Improving Learning in Low and Middle Income Countries?*