



HAL
open science

Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations en jeu

Nathalie Younes

► To cite this version:

Nathalie Younes. Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations en jeu. Sciences de l'Homme et Société. Université de Lorraine, 2020. tel-03032417

HAL Id: tel-03032417

<https://hal.science/tel-03032417v1>

Submitted on 30 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Lorraine – Nancy
Ecole doctorale : Sociétés, langages, temps, cultures



Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations en jeu

Mémoire de synthèse présenté le 6 octobre 2020 à Nancy

Nathalie Younès

Habilitation à diriger des recherches
Sciences de l'éducation (70^{ème} section du CNU)

Jury

Madame Emmanuelle Annot, Professeure des universités, Sciences de l'éducation,
Université de Rouen Normandie, CIRNEF – Rapporteuse

Monsieur Jean-Marie De Ketele, Professeur émérite
Université catholique de Louvain, Belgique – Rapporteur

Monsieur Jean-François Marcel, Professeur des universités, Sciences de l'éducation,
Toulouse, UMR EFTS – Rapporteur

Madame Christine Musselin, Directrice de recherche CNRS ; Centre de Sociologie des
Organisations, Sciences Po – CNRS – Présidente du jury

Monsieur Marc Romainville, Professeur des universités,
Université de Namur, Belgique – Examineur

Monsieur Saeed Paivandi, Professeur des universités, Sciences de l'éducation,
Université de Lorraine – Parrain scientifique

Sommaire

Préambule	4
Introduction	7
Partie 1 - Un positionnement épistémologique de l'évaluation dans les sciences de l'éducation et de la formation	14
1.1. Récurrence méthodologique et épistémologique de la pensée dichotomique	15
1.2. Une épistémologie de l'évaluation en chantier	18
1.2.1. Modèles de l'évaluation et théorisation	18
1.2.2. A partir des concepts de paradigme et d'incommensurabilité théorisés par Kuhn	20
1.2.3. Différentes cartographies de générations de modèles d'évaluation.....	23
1.3. Controverses autour des paradigmes en jeu	31
1.3.1. Lutte des paradigmes	32
1.3.2. Quatre principales controverses selon Guba et Lincoln	34
1.3.3. La prolifération des paradigmes avec Lather	38
1.3.4. L'hypothèse d'un paradigme synergique en évaluation	39
Partie 2 – Tensions au sein des dynamiques de l'évaluation : le cas de l'EEE	41
2.1. Tensions paradoxales	42
2.1.1. Entre validité de la mesure et validité conséquentielle	42
2.1.2. Entre démarche formative et outil de contrôle	43
2.1.3. Entre effets positifs, négatifs ou apparemment sans effet.....	47
2.1.4. L'impact des représentations de l'évaluation	49
2.1.5. Entre émotion, cognition et identité.....	51
2.2. Entre conceptions plurielles de l'enseignement, contextes et trajectoires	55
2.2.1. Typologies des conceptions de l'enseignement.....	58
2.2.2. Réflexivité critique et changements dans le contexte, facteurs de transformation.....	60
2.2.3. Enquête sur les dynamiques d'évolution pédagogique dans deux universités françaises .	61
2.2.4. Perspective élargie de l'EEE	69
Partie 3 – Les dispositifs d'évaluation comme médiations psycho-socio-techniques	72
3.1. La conception et l'opérationnalisation d'un dispositif d'EEE informatisé et l'étude de ses effets	73
3.1.1. A propos des dispositifs d'informatisation de l'évaluation.....	75
3.1.2. L'enjeu d'une articulation du quantitatif et du qualitatif pour l'appréhension des conditions de milieu et de sens.....	79

3.2. Evaluation interactive avec les boitiers de vote électroniques	81
3.2.1. Proposition d'un cadre intégrateur.....	82
3.2.2. Développement professionnel technopédagogique et usage des BVE	83
3.2.3. Dispositif de recherche/accompagnement de l'usage des BVE et développement professionnel pédagogique.....	86
Partie 4 - Dispositifs d'évaluation collaboratifs : transformations individuelles et collectives	91
4.1. Le tournant du partage dans l'EEE.....	92
4.1.1. Des figures de partage	93
4.1.2. Quelle mise en commun de l'EEE en France ?	95
4.1.3. Subjectivation et intersubjectivation dans un processus formatif	96
4.2 - Recherche collaborative sur les pratiques d'évaluation soutien d'apprentissage : l'évaluation subjectivante	98
4.2.1. Cadre théorique de l'enquête : subjectivation / intersubjectivation et évaluation soutien d'apprentissage	100
4.2.2. Les trois étapes de la recherche collaborative.....	102
4.2.3. Un pas vers une évaluation subjectivante et non seulement objectivante.....	109
4.3 - Co-construction d'un dispositif d'évaluation de l'enseignement à visée synergique.....	110
4.3.1. Une orientation subjectivante/intersubjectivante située du dispositif en jeu	110
4.3.2. Présentation du dispositif d'évaluation de l'enseignement dans un master de formation au métier d'enseignant	111
4.3.3. Trois études de cas.....	113
4.3.4. Types d'effets repérés dans les études de cas.....	121
4.3.5. Apports de la recherche et approfondissements requis.....	126
Conclusion.....	130
Références	133

Préambule

Une passion pour l'éducation, non exempte de formes de contestation, remonte à mon enfance et m'a orientée dans un cursus en psychologie sociale puis en sciences de l'éducation durant lequel j'ai privilégié un point de vue critique sur les formes instituées et une quête d'éducation ouverte et émancipatrice.

J'ai débuté mes études à Clermont-Ferrand à un moment où le professeur Jean Marc Monteil y développait un laboratoire de psychologie sociale. En élucidant l'impact des conditions de contexte sur les individus sans en omettre les relations de pouvoir, la psychologie sociale européenne des années 90, influencée par la sociologie bourdieusienne, démasquait en quelque sorte les processus à l'œuvre dans les injustices sociales et les destins individuels que l'école participe à fabriquer, sous couvert de méritocratie. Cette formation a profondément ancré en moi l'idée de la construction sociale du sujet humain et de sa sensibilité aux situations particulières dans lesquelles il s'inscrit. L'orientation du laboratoire était de fonder une psychologie sociale européenne d'orientation sociocognitive, cherchant en quelque sorte, à identifier les médiations sociales des processus cognitifs. Dans mon mémoire de DEA, j'ai tenté une construction théorique du concept de médiation socio-affectivo-cognitive, mobilisée dès le départ par une approche complexe tentant d'articuler différentes perspectives dans l'analyse des comportements et des phénomènes. L'approche expérimentale consistant à mettre les modèles théoriques à l'épreuve des faits et à imaginer des variations situationnelles pour comprendre les variations individuelles m'a fortement structurée, même si mes travaux se sont éloignés du paradigme positiviste de la méthodologie expérimentale mise en œuvre, étant autant intéressée par les singularités que par les tendances statistiques et désirant également agir dans un monde complexe et non artificiellement réduit à des variables isolées.

La prise de conscience de la surdétermination des situations évaluatives et de leur impact sur la construction du sujet, associée à la prise en compte de la longue histoire de la pensée de la pédagogie m'a sensibilisée au courant de l'évaluation formative comme « utopie porteuse » (Hajji, 1997), que la recherche se doit de documenter pour ne pas rester aveugle aux zones d'ombre de l'évaluation formative et établir les conditions de sa réalisation. La question des dispositifs de formation et d'évaluation informatisés a très vite été centrale dans mon expérience professionnelle en ingénierie de la formation et formation de formateurs, ainsi que dans le projet d'action qui a servi de support à ma thèse en sciences de l'éducation où il s'agissait de concevoir

un dispositif informatisé d'évaluation de l'enseignement par les étudiants. Dirigée par un spécialiste de la méthodologie de l'évaluation des dispositifs éducatifs, le professeur Gérard Figari, j'ai appris à articuler conceptualisation et opérationnalisation de l'évaluation. L'argument principal de la thèse est que l'évaluation de l'enseignement par les étudiants (EEE) constitue un passage critique accompagnant les fortes évolutions de l'université. L'enquête menée dans un établissement français a mis en évidence les effets paradoxaux de l'introduction d'un dispositif informatisé d'EEE à visée formative. En même temps qu'il a permis l'accès à de l'information pouvant conduire à des améliorations pédagogiques, il s'est heurté à de l'inertie et à des résistances, reflétant les contradictions de l'institution qui oscille vis à vis de l'EEE entre formalité administrative, visée formative et/ou visée de contrôle. Cette recherche a posé les bases d'une réflexion sur les conditions d'un effet formatif de l'EEE, notamment, sur l'importance de l'engagement de tous les acteurs dans un processus de régulation critique.

Maîtresse de conférences depuis 2008 à l'IUFM d'Auvergne devenu depuis ESPE (Ecole supérieure du professorat et de l'éducation) puis INSPE, j'ai continué à centrer mes recherches sur des questions relatives à l'évaluation, principalement au supérieur, mais aussi au secondaire et au primaire. Ces recherches se réalisent au sein du laboratoire ACTé (Activité, Connaissance, Transmission, éducation), dont l'orientation principale est « l'étude de l'activité professionnelle en situation, accordant une part importante à la subjectivité des acteurs étudiés tout en prenant en compte la contextualisation de leur activité ». Dans ce laboratoire interdisciplinaire, j'ai découvert les didactiques, l'éducation à la santé et l'ergonomie. D'autres collaborations interdisciplinaires et interprofessionnelles se sont enclenchées, d'autres ouvertures. Depuis 2016 je coordonne l'axe: Conception et évaluation d'outils et de dispositifs, dans lequel s'élaborent, au terme de recherches réalisées le plus souvent en collaboration avec les acteurs de terrain, des outils et dispositifs capables de jouer un rôle positif dans le développement professionnel de ces derniers, puis d'en évaluer la portée. C'est dans ce cadre que j'ai poursuivi mes recherches sur les dispositifs d'évaluation formative.

Les projets menés via les rencontres dans les réseaux de l'ADMEE (association de développement des méthodologies d'évaluation en éducation) dont j'assume la présidence depuis janvier 2017 et de l'AIPU (association internationale de pédagogie universitaire), sont constitutifs des évolutions de mon projet scientifique. Ainsi, dans le cadre du réseau de l'ADMEE RCPE (recherches collaboratives sur les pratiques évaluatives), j'ai orienté une partie de mes recherches sur les pratiques d'évaluation dans une approche collaborative, explorant progressivement plus particulièrement les processus évaluatifs à l'œuvre de manière formelle ou informelle dans ce type de recherches.

D'autre part, les collaborations avec Saeed Paivandi rencontré lors des colloques de l'AIPU m'ont initiée à la perspective expérientielle de la sociologie interactionniste qui se révèle une influence majeure, en écho avec mon ancrage initial en psychologie sociale et avec l'approfondissement, dans mes travaux, des dynamiques collaboratives et de la culture universitaire.

Enfin, les échanges au long cours avec une mère philosophe et un goût ancien pour cette discipline et ses ouvertures intellectuelles et existentielles ont été déterminants. Dès mes premières années d'étude en psychologie, j'ai opté pour des unités d'enseignement libres en philosophie, épistémologie et esthétique ; orientations qui continuent à nourrir mes travaux.

Introduction

Ce projet scientifique s'inscrit dans un parcours de recherche construit autour de l'évaluation de/dans l'enseignement. Omniprésente et controversée, partagée et polysémique, implicite et explicite, intuitive et rationalisée, intrusive et privée, dévastatrice et constructive, pragmatique et utopique, l'évaluation offre un espace d'analyse et d'intervention particulièrement stratégique pour qui s'intéresse, sur un plan aussi bien théorique que praxéologique, à l'enseignement et à la formation. Car l'indéfinition du terme, la multiplicité de ses formes, l'étendue des débats et controverses à son propos, la variété des manières de faire et des impacts ouvrent le champ des possibles en recherche et sur le terrain. Bien que souvent impensée et réduite dans le champ des pratiques, l'évaluation oriente pourtant profondément, de façon directe et indirecte, les processus d'enseignement et d'apprentissage, et plus largement les construits sociaux et individuels. Il s'agit d'une pratique sociale engagée dans des champs culturels et normatifs, d'une activité située interpersonnelle. Son approche scientifique s'inscrit ainsi à la fois dans une orientation socio-affectivo-cognitive, complexe et contradictoire et dans des dispositifs techniques.

Notre cadre de recherche théorique et méthodologique, qui se veut ouvert aux apports de différentes disciplines et aux contradictions inter et intra disciplinaires, porte plus particulièrement sur les liens entre les dispositifs d'évaluation de l'enseignement et les transformations en jeu. Son élaboration s'est faite progressivement à travers un parcours jalonné de rencontres avec des figures marquantes du champ pédagogique : durant une formation initiale en psychologie puis une thèse en sciences de l'éducation et par la suite dans de nombreuses collaborations scientifiques interculturelles. Chaque rencontre a été un moment et un nœud dans l'explicitation intellectuelle et expérientielle d'un univers conceptuel des questions d'évaluation en éducation et formation. Les recherches sur l'évaluation de l'enseignement par les étudiants (EEE), amorcées en thèse, ont été poursuivies et élargies aux questions d'évaluation des apprentissages, et plus généralement aux transformations de la pédagogie, étant données leurs intrications. Les pistes successivement ou concomitamment suivies ne sont pas des voies séparées. Les repères et phénomènes étudiés dans le champ de l'évaluation de l'enseignement par les étudiants et dans celui de l'évaluation des apprentissages qui ont des proximités en termes de référents et de méthodologies, se trouvent indissociables des évolutions des conceptions de l'enseignement en leurs ontologies et épistémologies, plus ou moins explicitées. L'étude

cartographique des positionnements épistémologiques de l'évaluation dans les sciences de l'éducation et de la formation conduit à souligner la pluralité. Il est clair que des antagonismes sont plus ou moins affirmés et que leur coexistence est souvent difficile.

Parmi les soubassements déterminants, la « référentialisation » ou méthodologie de l'évaluation partagée (Figari, 1994 ; 2001 ; 2014) comme forme d'enquête évaluative amorce l'importance de trouver des terrains d'entente situés, ce qui a pris de plus en plus de place dans les orientations des travaux que je développe. La référentialisation détermine des référents à la fois dans le champ théorique et dans le contexte où elle se déroule en impliquant les parties prenantes. Elle conduit aussi à construire des instruments d'analyse dans lesquels les critères sont justifiés et des indicateurs élaborés pour en vérifier le degré de réalisation. Sur la base d'échanges préalables, de réflexions et de négociations, les partenaires peuvent se mettre d'accord sur le recueil et l'interprétation des informations ainsi que sur les actions consécutives à mettre en œuvre. Un tel dispositif tend à prendre en compte la constitution d'un milieu *in situ*, dans lequel les sujets échangent par rapport à la détermination des objectifs, des procédures et des stratégies à privilégier.

Vers une perspective d'écologie des pratiques de l'évaluation de l'enseignement

Toutefois de tels objectifs restent difficiles à réaliser. C'est ainsi que j'en suis venue à étudier de manière critique cette méthodologie de la référentialisation et à renforcer une perspective écologique de l'évaluation dans l'enseignement, à savoir prenant en compte davantage les caractéristiques et dynamiques propres à chaque milieu, considéré en termes d'environnement d'études et de cultures universitaires tant disciplinaires que pédagogiques mais aussi en tant que dynamiques subjectivantes/intersubjectivantes étant donné l'importance du sujet et de ses interrelations dans le processus de formation. Loin de renvoyer à des replis subjectivistes dans une orientation qui serait seulement psychologique, ce qui est en jeu c'est le dépassement de la séparation entre réalité objective et réalité subjective afin d'être à même d'appréhender une « écologie des pratiques » (Stengers, 2006). Dans cette perspective, la subjectivation n'est pas isolée mais considérée dans les interactions avec les autres et avec le contexte social.

Dans un « modèle écologique des pratiques enseignantes » (Marcel 2009, p. 48) les pratiques enseignantes sont envisagées comme un système en trois composantes interdépendantes : l'acteur, le contexte et l'action située. L'acteur est envisagé au travers de caractéristiques socioprofessionnelles, d'une histoire personnelle et professionnelle, de valeurs, d'expérience, d'affects et sensibilité, mais aussi de ressources. Le contexte renvoie à des dimensions tant relationnelles entre les différentes parties prenantes (élèves, autres enseignants, partenaires), que matérielles (la classe, les ressources pédagogiques, etc.), institutionnelles (textes officiels,

programmes, etc.), symboliques ou culturelles. Quant à l'action en situation, elle est « marquée par une contingence irréductible ».

Une perspective d'écologie des pratiques d'évaluation engage une approche holistique qui tente d'articuler différents niveaux d'environnement et de sens (Younès, 2013).

Subjectivation/intersubjectivation

Dans cette perspective d'écologie des pratiques d'évaluation, nous avons été amenée à mettre l'accent sur la question des interrelations entre construction de soi et construction de l'autre qui est largement présente dans différentes disciplines des sciences de l'humain et de la société.

Ainsi, elle est au cœur de la fondation phénoménologique d'une sociologie qui envisage la réalité comme le produit de l'action intersubjective (Schütz, 1987), comme de la psychologie sociale, définie comme la « science des interactions avec autrui » (Bègue et Desrichard, 2013), voire la science des conflits avec autrui (Moscovici, 1984 p.6) et qui envisage autrui au cœur d'un soi en construction permanente (Mead, 1963, Martinot, 2016). Mais elle prend des tournures très différentes selon les paradigmes de référence. Peu traitées comme telles dans les approches positivistes des SHS, les notions de subjectivation et d'intersubjectivation, centrales dans notre approche de l'évaluation, requièrent des précisions conceptuelles autour des notions de sujet, intersubjectivité, subjectivation et intersubjectivation.

En caractérisant l'évaluation comme un travail du sujet, Vial exprime l'idée que « l'évaluation se caractérise par le rapport que les sujets entretiennent avec la valeur, avec ce qui importe, avec ce qui fait sens pour eux » (Vial, 2012, p. 405). Qui est le sujet de l'évaluation (évaluateur et évalué) ? Penser le sujet de l'évaluation c'est penser d'emblée les relations entre le sujet connaissant, le sujet de l'expérience et le sujet de l'action plongé dans le monde. La relation aux autres est constitutive de l'expérience humaine, ce qui fait de l'intersubjectivité la condition de la subjectivité.

L'approche rationaliste du sujet ancrée dans la philosophie cartésienne et kantienne qui se poursuit dans les orientations cognitivistes et neuroscientifiques met l'accent sur la connaissance et les représentations. Quant au sujet de l'action, il est aussi souvent pensé comme libre et capable. Selon Touraine (1994, p.23), il est « l'effort de transformation d'une situation vécue en action libre ». Accroître cette « capacité » serait étendre la possibilité d'être sujet. Dans la perspective phénoménologique initiée par Husserl, le sujet c'est celui qui donne sens à l'expérience (Husserl, 1947), l'accent est mis sur la conscience comme puissance organisatrice d'une expérience dans une tension entre l'être, le vivre et le faire. Le sujet se projette sur le monde, l'invente et s'invente lui-même. Il est constamment transformé à partir de cette expérience dans laquelle il est engagé et qu'il réalise mais qui le fait aussi.

Dans ces trois approches, le focus est porté sur le sujet, connaissant, éprouvant, agissant. C'est à travers ce focus que l'intersubjectivité est appréhendée. Initialement forgé par Kant, le concept d'intersubjectivité renvoie dès l'origine à l'idée que les hommes sont des sujets pensants capables de prendre en considération la pensée d'autrui dans leur propre jugement. Dans la perspective phénoménologique initiée par Husserl, l'accent a été mis sur l'expérience de l'autre à partir de la perception et de l'empathie, décrivant des formes de coprésence au monde. Il développe l'idée d'une décentration du sujet pouvant s'imaginer à la place de l'autre dans une forme de réversibilité des points de vue. Un autre courant trouve son origine dans les travaux de Fichte qui, s'intéressant aux interactions entre sujets, considère l'intersubjectivité dans sa dimension relationnelle (Aubert, 2008). La relation aux autres, constitutive de l'expérience humaine dans son rapport à soi et au monde fait de l'intersubjectivité la condition de la subjectivité, et de la reconnaissance d'autrui, la condition de la conscience de soi (Fischbach, 2000, p 143).

Une telle lecture philosophique du sujet entre en résonance avec les travaux ultérieurs conduits en psychologie qui vont opérationnaliser l'étude. Le sujet attribue des significations en fonction de représentations structurant la perception subjective à travers un processus de sélection de l'information. « Simultanément instrument de connaissance et instrument d'action, elle [la représentation] est l'occasion pour le sujet d'organiser les informations reçues de son milieu afin de les rendre intelligibles et, par-là, de leur imprimer un contenu dynamique, permettant d'orienter sa conduite. Ainsi, par la représentation, le sujet donne un sens au monde et y trouve des points d'appui pour pouvoir agir » (Moscovici et Paicheler, 1973, p.40). Mais ces représentations sont elles-mêmes aussi des produits sociaux – normes et valeurs – des connaissances socialement élaborées et partagées (Jodelet, 1984, p. 366). Le social est aussi incorporé dans l'individu. Si selon Husserl (1996, p. 76) « la subjectivité est le premier en soi », cette construction subjective du monde de la vie, n'est pas indépendante des formes sociales préexistantes qui déterminent les visions du monde des personnes et des groupes, les regards portés par et sur les autres par et sur soi, ni des formes sociales émergentes des interactions dans les situations sociales. Il s'agit d'une dynamique de création continue dans la relation à un milieu. Ainsi selon (Mead, 1934/1963 ; p. 141) : « Quand nous atteignons le soi, nous touchons à une certaine sorte de conduite, un certain type de processus social qui comprend l'interaction de différents individus, et qui implique des individus engagés dans des activités coopératives. C'est dans ce processus que peut se développer un soi comme tel. » Via les interactions avec l'« autrui généralisé » (qui comprend les autres mais aussi soi-même et les objets), ce processus est un processus de transformations mutuelles permanent. Mead et ses continuateurs ont particulièrement théorisé ce processus de co-création de l'esprit et du soi, co-définis comme

résultats dynamiques de la relation organisme-monde en constante transformation (Joas, 1999 ; Brassac, 2005).

Comment accéder au sens de l'action, éminemment subjectif, flou et mouvant ? C'est tout le projet intersubjectif de la sociologie weberienne reprise par Schütz (1987) et son apport épistémologique. Accessible seulement au sujet agissant, dans des configurations floues et mouvantes, ce sens renvoie à la fois à un devenir, un projet et à la conscience réflexive sur un acte passé. Selon Schütz: « Les objets de pensée construits par le chercheur en sciences sociales afin de saisir la réalité sociale, doivent être fondés sur des objets de pensée construits par le sens commun des hommes vivant quotidiennement dans le monde social. De la sorte, les constructions des sciences sociales sont, pour ainsi dire, des constructions du second degré, c'est-à-dire des constructions de constructions faites par les acteurs sur la scène sociale, dont le chercheur doit observer le comportement et l'expliquer selon les règles procédurales de sa science » (Schütz, 1962. p. 59). Ces constructions au second degré sont des constructions objectives idéalisées et typiques. Sur la base de l'interprétation subjective, le chercheur élabore la construction d'un comportement typique pour tenter de dégager des invariants (Laoureux et Frère, 2013, p.4).

Étudier les processus de subjectivation et d'intersubjectivation en jeu dans l'évaluation ne se limite pas à l'étude des interactions et à l'idée d'intersubjectivité puisqu'il s'agit de considérer des processus en construction. Nous définissons la subjectivation comme le processus par lequel le sujet se recrée en fonction de strates d'expériences acquises au fil du temps, d'une adaptation aux milieux et d'un projet en devenir.

L'évaluation est un élément déterminant de ce processus. Qu'il s'agisse du regard de l'autre ou du regard de soi, indissociablement liés, au bout du compte, c'est la valeur et la reconnaissance qui sont en jeu. C'est ce qui nous a amenée à proposer le concept d'évaluation subjectivante, par opposition à une évaluation objectivante de l'être de l'évalué, en le positionnant de l'extérieur au regard de critères normatifs, dont les travaux de psychologie sociale ont bien montré le coût humain (Butera, Buchs et Darnon, 2011, Butera et al., 2020). Dans la filiation de l'évaluation formative, l'évaluation subjectivante considère l'évalué comme un sujet capable de se situer et de s'impliquer dans une dynamique de progression. Ce type d'évaluation, par l'importance accordée aux dynamiques de subjectivation et d'intersubjectivation s'éloigne de toute tentation positiviste concevant l'évaluation comme un système seulement objectivant. Les « strates de la subjectivation » qui s'élaborent autour de la création d'une dynamique d'évaluation plus collaborative quant au sens, cadre, objectifs et critères de l'évaluation et par rapport à son inscription dans une dynamique de progression (Younès, Sasse et Darj, 2016).

Médiations psycho-socio-techniques

Ces considérations m'ont ainsi amenée au fil du temps, depuis la thèse jusqu'à mes travaux actuels, à approfondir la problématique des dispositifs d'évaluation comme médiations psycho-socio-techniques en leurs conditions de milieu et de sens, impliquant d'intégrer les dimensions culturelles variables suivant les spécificités locales et les circonstances. Pensés à partir des sciences de l'éducation mais aussi à partir des travaux philosophiques notamment de Foucault (2001)¹ et d'Agamben (2007), de tels dispositifs sont entendus dans leur dimension technologique (Albero, 2010, a et b) dans la manière d'établir et d'instituer des agencements et des contraintes, et dans leur dimension politique de pouvoir. Ils sont entendus aussi dans leurs dimensions subjectivantes/intersubjectivantes. La puissance des dispositifs sur la formation et la transformation des personnes peut s'avérer émancipatrice ou aliénante, étant donné l'emprise qu'ils exercent, comme souligné par Agamben (2007, p.31) : « J'appelle dispositif tout ce qui a d'une manière ou d'une autre la capacité de capturer, d'orienter, de déterminer, d'intercepter, de modeler, de contrôler et d'assurer les gestes, les conduites, les opinions et les discours des êtres vivants ». Ce sont des processus de désubjectivation qui sont à l'œuvre.

La référence à Dewey (1993, 2010, 2011) est également mobilisée à propos de son approche des sciences sociales comme dispositif intégrant enquête et politique. Selon cette démarche, tout concept ou principe est mis à l'épreuve en tant qu'outil d'enquête. De même, chaque mesure politique est à considérer comme une hypothèse de travail à étudier et non comme un programme à appliquer. Des observations fines et continues des conséquences d'une mesure politique et sa révision en fonction des analyses menées en fondent le caractère expérimental : « Si ces deux conditions étaient remplies, les sciences sociales pourraient former un dispositif capable de conduire les investigations, d'enregistrer et interpréter (organiser) leurs résultats. Dans ce cas, on ne considérerait plus ce dispositif comme la connaissance elle-même, mais comme l'ensemble des moyens intellectuels destinés à découvrir les phénomènes ayant une portée sociale et à comprendre leurs significations » (Dewey, 2010, p. 305).

L'ensemble des recherches présentées dans les quatre parties du projet scientifique d'HDR, qui resitue comment les dispositifs entendus comme médiations psycho-socio-techniques et les transformations se reliant m'ont conduit ainsi à développer l'hypothèse d'un paradigme synergique de l'évaluation de l'enseignement.

¹ Foucault définit le dispositif comme « un ensemble résolument hétérogène comportant des discours, des institutions, des aménagements architecturaux, des décisions réglementaires, des lois, des mesures administratives, des énoncés scientifiques, des propositions philosophiques, morales, philanthropiques ; bref, du dit aussi bien que du non-dit » (2001, p.299).

Dans une première partie, nous envisageons comment l'évaluation se positionne épistémologiquement dans les sciences de l'éducation et de la formation, selon différents paradigmes et en termes de générations d'évaluation. La deuxième partie expose les tensions et les dynamiques de changement dans le cas de l'évaluation de l'enseignement par les étudiants. Dans une troisième partie, nous traitons des dispositifs d'évaluation comme médiations psychosocio-techniques dans le cadre d'évaluations avec le numérique. La quatrième partie considère les relations entre dispositifs d'évaluation collaboratifs et transformations individuelles et collectives qui impliquent de mettre en synergie les différents éléments constitutifs.

Partie 1 – Un positionnement épistémologique de l'évaluation dans les sciences de l'éducation et de la formation

Selon Gaston Mialaret, le fondateur de la discipline en France, « Les sciences de l'éducation sont constituées par l'ensemble des disciplines qui étudient, dans des perspectives différentes mais complémentaires et coordonnées, les conditions d'existence, de fonctionnement et d'évolution des situations et des faits d'éducation » (1976, p. 32). Cette orientation épistémologique a ensuite dirigé la deuxième génération d'universitaires qui a pensé les sciences de l'éducation à partir des sciences humaines et sociales « mères » : on est psychologue, sociologue, historien, anthropologue dans la discipline universitaire des Sciences de l'éducation et on fait de la recherche sur des objets éducatifs qui nécessitent la spécificité de l'approche plurielle, pluridisciplinaire propre des sciences de l'éducation et fondant progressivement un regard propre (Altet, 2016, p.28). Mais il n'est pas si simple de se définir comme enseignant-chercheur en sciences de l'éducation quand on ne se réfère pas à une autre discipline dont on partage alors un cadre théorique et méthodologique spécifique. Les différentes appellations selon les pays, tout en étant proches, manifestent la complexité de ce champ. Que l'on s'inscrive dans la science de l'éducation (cas de l'Allemagne), l'éducation (cas du Brésil ou du Japon par exemple) ou les sciences de l'éducation (cas de la France), une discipline universitaire définit un champ de connaissances et pas seulement un champ de pratiques et de politiques sur lesquelles les autres disciplines des Sciences Humaines et Sociales produiraient des connaissances. C'est la position que je défends à l'instar de Charlot notamment, pour qui il s'agit « d'un champ de savoirs fondamentalement métissé où se croisent, s'interpellent et parfois se fécondent, d'une part des connaissances, concepts et méthodes originaires de champs disciplinaires multiples, d'autre part des savoirs, des pratiques, des finalités éthiques et politiques » (Charlot, 2008, p.4). Il assume alors ce positionnement d'une « discipline épistémologiquement faible : mal définie, aux frontières floues, aux concepts fluides, qui n'a pas et n'aura jamais l'apparente pureté et clarté de la psychologie et de la sociologie » Comme lui, j'ai relevé le pari d'une spécificité des sciences de l'éducation et inscrit cette spécificité dans mes recherches. Pour autant peut-on dire que le « métissage », le « flou » et la « fluidité » fassent des sciences de l'éducation une discipline « épistémologiquement faible » ? Epistémologiquement inconfortable et éclatée, à coup sûr, mais ce sont peut-être justement ces caractéristiques qui peuvent constituer une force

épistémologique pour la compréhension et l'explication des phénomènes éducatifs, complexes par nature. Je ne pense pas non plus que la psychologie ou la sociologie présentent ce caractère « d'apparente pureté et clarté » qu'il invoque car ce serait méconnaître la virulence des débats qui font rage dans ces disciplines et ce depuis leurs origines.

1.1. Récurrence méthodologique et épistémologique de la pensée dichotomique

Il est certain que les sciences de l'éducation comme les sciences humaines et sociales en général ne sont pas épargnées par une pensée dualiste et clôture disciplinaire qui semblent liées et que l'on trouve aussi bien sur le plan méthodologique qu'épistémologique. Figari (2014) remarque ainsi que la littérature sur l'évaluation est souvent construite autour de couples d'opposition : objectiviste /subjectiviste, qualitative/quantitative, formative/certificative, individuelle/collaborative, etc.

Cette pensée dichotomique est également à l'œuvre en sociologie si l'on suit les analyses de Berthelot (1991, p.63), pour qui « parler de méthode en sociologie équivaut très souvent à mobiliser une série d'oppositions récurrentes : quantitatif/qualitatif ; longitudinal/transversal ; dynamique/statique ; intensif/extensif ; objectivisme/subjectivisme ... ».

Richardot (2006), à partir d'une enquête par entretiens conduite dans les années 2000 auprès d'une dizaine de chercheurs clés de la discipline, montre que la psychologie sociale semble s'organiser autour d'une opposition structurante dans laquelle le champ social serait systématiquement opposé au champ scientifique. La recherche scientifique serait caractérisée par son caractère fondamental, la neutralité du chercheur, une visée explicative, des études conduites en laboratoire et des méthodes quantitatives. Par opposition, la recherche appliquée, à visée descriptive, moins voire peu scientifique, serait conduite sur le terrain, par un chercheur impliqué et des méthodes qualitatives.

On voit que derrière les oppositions sur les méthodes, d'autres dimensions sont en jeu, dont on peut se demander si elles sont d'ordre épistémologique, idéologique ou positionnel. Dans le champ de la sociologie, Lapeyronnie (2004) qualifie d'« académisme radical », le positionnement scientifique de sociologues, qui de leur point de vue « surplombant », au nom d'une « neutralité axiologique » « échapperaient aux déterminismes sociaux qu'ils pourraient repérer chez les autres », un positionnement selon lui plus idéologique que scientifique. De manière réciproque, les recherches sociologiques s'inscrivant dans les courants de la sociologie de l'acteur sont qualifiées par d'autres, de constructions « semi-savantes » qui « s'enorgueillissent d'être en symbiose avec le "point de vue" des acteurs » produites par les « sociologues-journalistes »

(Bourdieu et Wacquant, 1998, p.116). On peut noter, en remontant aux origines de la discipline, que Simmel qui ne séparait pas sociologie, psychologie et philosophie et qui fut un inspirateur important de l'école sociologique de Chicago, fut récusé comme sociologue par Durkheim dont il était le contemporain (Lepenies, 1990). Dans le champ de la psychologie sociale également, entre psychologie fondamentale et psychologie appliquée, le débat fait rage quant à savoir qui fait la recherche la plus scientifique, au regard des objectifs, modalités de construction des savoirs et méthodologies des recherches. Pour certains, la psychologie sociale se définirait par sa méthode. L'expérimentation en laboratoire (ou la quasi expérimentation sur le terrain) serait ainsi la seule garantie de la production des connaissances scientifiques et le terrain serait le lieu d'application des connaissances produites en laboratoire (Beauvois, 2001 ; Py, Somat et Baillé, 1999). Mais pour d'autres (Moser, 2006), la psychologie sociale appliquée est également à même de produire de la connaissance scientifique comme en témoigne notamment l'existence de revues scientifiques dans ce champ².

Dans le domaine de l'évaluation en éducation et formation, alors que les partisans d'une « evidence based education » contestent le sérieux, la rigueur et pour tout dire la scientificité des études évaluatives qualitatives, d'autres critiquent avec vigueur les pratiques actuelles d'évaluation quantitative qui se développent jusqu'à devenir les seules pratiques légitimes. En ce qui concerne l'évaluation de la recherche, ces critiques peuvent provenir des experts de la mesure eux-mêmes et se situer sur un plan technique en montrant « le mauvais usage de faux indicateurs » (Gingras, 2008). Elles peuvent également porter plus fondamentalement sur le modèle de société et la place des SHS dans ce modèle qu'elles portent (Gori et Del Volgo, 2009) ainsi que sur les dérives pouvant être liées aux indicateurs (Goasdoué, 2013 ; Vinokur, 2006).

Ce texte n'est pas le lieu pour chercher les raisons qui amènent à produire des conduites de distinction et d'exclusion « à partir de normes scientifiques parfois radicalisées jusqu'à la caricature » pour reprendre les mots de Lordon (1997, p. 29) dans le champ de l'économie cette fois. Mais prenant acte du fait qu'on les retrouve dans l'ensemble du champ des sciences humaines et sociales, il nous faut quand même nous interroger sur leur sens et leur portée. Les oppositions méthodologiques sont classiquement référées au niveau épistémologique : Pour faire vite, la nécessaire réduction des sciences humaines et sociales aux sciences de la nature selon le modèle de Durkheim (traiter les faits sociaux comme des choses) versus la reconnaissance de leur spécificité comme sciences anthroposociales dans la tradition weberienne. Mais plusieurs arguments permettent de soutenir que ce dualisme dichotomique « est de niveau pragmatique et

² Comme par exemple : Applied Psychology, Journal of Applied Psychology

que, bien au-delà des difficultés épistémologiques sur lesquelles il peut prendre appui, il fonctionne comme norme et référence obligée au sein d'un champ d'énonciation déterminé » (Berthelot, 1991, p.61). Derrière l'étendard de la scientificité il y a des questions d'identité et de pouvoir. Ce constat n'est pas neuf. La culture légitime, c'est toujours la sienne.

Ayant navigué dans différents champs et constaté les logiques d'exclusion qui prévalent dans chacun, nous ne pouvons que constater la permanence de l'analyse de Bourdieu (1976, p.88) selon laquelle : « Le champ scientifique est un champ social comme un autre avec ses rapports de forces et ses monopoles, ses luttes et ses stratégies, ses intérêts et ses profits, qui a pour enjeu spécifique la lutte pour le monopole de l'autorité scientifique (prestige, reconnaissance, célébrité, etc.) ». En effet il convient de ne pas être dupes de cette réalité qui, tout en s'exprimant de manière violente et récurrente, ne saurait pour autant être réduite seulement à ces rapports de force, sans quoi il nous faudrait admettre que l'activité scientifique ne se distinguerait absolument pas des autres activités du champ social. Cependant, il importe également de reconnaître que l'activité scientifique comme activité humaine est éminemment personnelle et sociale. La connaissance n'est pas construite de manière détachée, elle engage la personne du scientifique dans tout son être affectif et cognitif, bien loin du détachement impersonnel qui est donné à voir dans le formalisme de la communication scientifique standard (Polyani, 1958). En prenant plusieurs exemples, Feldman (2002) montre l'importance des dimensions affectives et spirituelles dans les rapports des chercheurs à la science.

La culture disciplinaire implique des positionnements épistémologiques sans lesquels la pratique scientifique n'est pas possible. Mais l'absence de méta-connaissance dans ce domaine amène souvent les praticiens de la recherche à réduire les fondements de leurs approches, en ne percevant que des bribes et en entretenant d'une certaine manière une pensée dualiste. Il ne s'agit pas de nier la réalité et l'efficacité discursive voire taxonomique des oppositions évoquées mais d'interroger leur statut épistémologique et logique au regard notamment des modèles impliqués et de la notion d'incommensurabilité des paradigmes introduite par Kuhn et reprise par Feyerabend. Selon cette proposition, une science ou un état de science fonctionne comme une langue dont la syntaxe et le vocabulaire définissent simultanément les énoncés possibles et les objets auxquels ils s'appliquent : entre deux formations de ce type rien n'est commun, seules de traductions approximatives sont possibles. Ce qui nous amène à nous centrer sur l'épistémologie des modèles de l'évaluation, qui donne lieu à de fortes controverses. Même si la critique de la pensée dualiste est transversale, ces débats évoquent les querelles dans les différentes disciplines des SHS mettant en évidence l'importance des modèles engagés dans l'épistémologie de l'évaluation.

1.2. Une épistémologie de l'évaluation en chantier

La question des modèles de l'évaluation s'est fortement posée dans les années 80 avec l'émergence de la professionnalisation des évaluateurs mais aussi les évolutions sociétales, scientifiques et épistémologiques.

Questionner les modèles et modélisations de l'évaluation en éducation dans une perspective épistémologique comme se sont proposés de le faire Mottier Lopez et Figari (2012) n'est certainement pas une chose aisée tant la diversité des modèles est grande, contribuant suivant les auteurs à un « brouillage » du champ de l'évaluation (p.10), à une pensée éclatée (p15). Dès 1983, Chen et Rossi ont souligné l'importance d'une « évaluation dirigée par la théorie ». Madaus, Stufflebeam et Scriven (1989) comme Madaus, Stufflebeam, et Kellaghan, (2000), considèrent que les évaluateurs doivent être capables de conceptualiser l'évaluation, c'est-à-dire non seulement être au clair avec les concepts centraux du champ (évaluation, valeur, mérite, prix, critère, objectif, standard, besoin, validité, fidélité, amélioration, formatif, sommatif...) mais aussi pouvoir les intégrer dans un système cohérent. Figari (2012, p. 212) nous engage à avancer sur la voie d'une théorisation de l'évaluation consistant « à identifier et schématiser les dimensions épistémologiques de l'évaluation en éducation telle qu'elle est pensée et non pas seulement pratiquée aujourd'hui ». La même année, Vial proposait un ouvrage pour « se repérer dans les modèles de l'évaluation », considérant le modèle comme « la matrice des dispositifs » mais la modélisation comme un travail du sujet (Vial, 2012, p.13). Qu'elle soit pratiquée par des professionnels de l'enseignement ou de la recherche, l'évaluation se réalise dans le champ social en fonction de la définition qui lui est donnée, elle-même dépendante des épistémologies plus ou moins profanes et plus ou moins implicites (Vial, 2013) que les professionnels portent. D'après cet auteur, il est donc « vain de chercher la définition ultime de l'évaluation, chaque modèle en donnant une version, il ne reste que la signification épistémologique : l'évaluation est le travail des valeurs » (p.14).

1.2.1. Modèles de l'évaluation et théorisation

Dans ces différentes démarches, l'enjeu n'est pas seulement d'identifier les modèles de l'évaluation mais de « s'interroger sur les « modèles d'intelligibilité » que l'évaluation convoque et leur « modélisation en tant que construction de systèmes symboliques et représentationnels par lesquels elle se définit et s'opérationnalise dans un contexte historique et social donné » (Mottier Lopez et Figari, 2012, p.7). La diversité des approches de l'évaluation renvoie à différents arrière-plans théoriques, expériences et visions du monde qui façonnent des orientations philosophiques particulières et des préférences méthodologiques et pratiques qui conduisent les professionnels à

proposer des méthodes parfois radicalement différentes pour la conduite de l'évaluation, la collecte et l'interprétation des données. Ces différences dans les approches de l'évaluation trouvent leurs origines dans des présupposés ontologiques et épistémologiques renvoyant à la nature de la connaissance (épistémologie) et à la nature du réel (ontologie) en fonction de contextes particuliers. Car en effet, si une histoire de l'évaluation en éducation et formation reste à écrire, les éléments dont on dispose montrent bien comment la pratique de l'évaluation (et la pensée sur la pratique de l'évaluation) est liée aux conditions sociales et historiques dans lesquelles elle s'inscrit. Selon Barbier (1983), il y aurait une histoire en parallèle, bien plus récente, de l'évaluation des personnes et de celle des dispositifs et des actions de formation. Il analyse comment l'évaluation des personnes s'est développée comme une rationalisation croissante des moyens de gestion sociale des caractéristiques des individus sur le marché du travail à travers le repérage des différences individuelles, professionnelles et scolaires. Les pratiques d'évaluation des dispositifs et actions de formation, nées dans un cadre industriel et généralisées à l'ensemble des secteurs de la formation, se comprennent également en fonction de logiques de contrôle et de gestion.

Les analyses de Madaus, Stufflebeam et Scriven (1989) reprises par Madaus, Stufflebeam, et Kellaghan, (2000), mettent en perspective les développements de l'évaluation au regard de dimensions sociopolitiques mais aussi technoscientifiques. Ils identifient 6 périodes ou âges de l'évaluation, 7 dans la réédition de leur ouvrage.³ Dans cette analyse, on voit comment

-
- ³L'Âge de réforme (fin du 18e siècle – 1900) serait caractérisé par l'invention des notes pour estimer la performance des élèves puis la mise au point de tests pour évaluer aussi la performance des méthodes d'enseignement.
 - L'Âge d'efficacité et des tests (1900-1930), articulé au modèle de l'organisation scientifique du travail (Taylorisme). La mise au point de tests objectifs était alors considérée cruciale pour déterminer la qualité de l'enseignement. Évaluation et mesure étaient considérées comme synonymes.
 - L'Âge tylerien (1930-1946). Tyler a posé les bases du modèle par objectif : définir les objectifs en termes de comportements mesurables pour évaluer l'efficacité de l'enseignement en comparant les comportements avec les objectifs.
 - L'Âge de l'innocence (1946-1957) ainsi dénommé car correspondant à cette période de croissance et d'optimisme de la société américaine. Pendant cette période les travaux de Bloom et ses collaborateurs ont approfondi les acquis de Tyler et permis de proposer une taxonomie des objectifs d'apprentissage en termes de domaines cognitifs, taxonomie encore utilisée de nos jours en formation.

l'expansion et la technicisation de l'évaluation sont indissociables de l'expansion de la société industrielle et positiviste, dans laquelle elle devient une technique de gouvernement au regard d'une certaine idée de la science.

La perspective est également historique mais plus approfondie sur le plan épistémologique dans les propositions de Shadish, Cook et Leviton (1991), qui analysent les « théories » de l'évaluation produites aux Etats-Unis entre les années 60 et les années 90 selon un modèle de la théorie défini en cinq composantes. Ils insistent sur le rôle déterminant d'une théorie de l'évaluation : si l'évaluation des programmes sociaux est un champ conduit par la pratique, la théorie est fondamentale en ce qu'elle développe les raisons théoriques des prescriptions pratiques. Ils déplorent que de nombreux ouvrages sur l'évaluation soient des listes a-théoriques de méthodes. Pour eux, le but fondamental d'une théorie de l'évaluation est de spécifier des pratiques faisables par les évaluateurs pour construire une connaissance de la valeur des programmes qui puisse être utilisée pour améliorer la résolution des problèmes sociaux pour lesquels les programmes sont pertinents. Suivant cette définition, une théorie de l'évaluation doit explicitement décrire et justifier les 5 composantes suivantes : Une théorisation de l'objet (en l'occurrence le programme social) ; une théorisation de la connaissance produite par l'évaluation, autour d'une ontologie, d'une épistémologie et d'une méthodologie ; une théorisation de l'utilisation de la connaissance produite par l'évaluation, une théorisation de l'attribution de la valeur, une modélisation de la pratique de l'évaluation (Shadish, Cook et Leviton, 1991, p.36).

1.2.2. A partir des concepts de paradigme et d'incommensurabilité théorisés par Kuhn

Quand il est question de l'épistémologie de l'évaluation (et d'épistémologie en général), de nombreux auteurs, tels notamment De Ketele (1993 et 2019), Dubois et Marceau (2005), Mottier Lopez et Figari (2012) se réfèrent à Kuhn et sa théorie des révolutions scientifiques initialement publiée en 1962, marquée par le concept de paradigme.

-
- L'Âge de l'expansion (1958-1972). En pleine guerre froide, l'éducation s'est vue attribuer un rôle crucial au regard des objectifs nationaux. Il s'agissait de mesurer la réussite des nouveaux curriculums en mathématiques, sciences et langues étrangères. Les modèles de mesure progressent avec l'apparition des tests critériés.
 - L'Âge de la professionnalisation (1973-1983). Emergence de la profession d'évaluateur et de programmes de formation universitaire à l'évaluation.
 - L'Âge de l'expansion et de l'intégration (1983-années 2000) à l'heure de la réduction des ressources des programmes éducatifs et de l'inscription de l'évaluation dans une approche économique.

Dans la théorie de Kuhn, la notion de paradigme est effectivement centrale comme une sorte de noyau organisateur au fondement du consensus entre des membres d'une communauté scientifique spécifique quant au choix légitime des problèmes concrets à résoudre, aux méthodologies à utiliser et aux manières de trouver des solutions concrètes. Un paradigme est issu "d'une découverte scientifique universellement reconnue qui, pour un temps, fournit à la communauté de chercheurs des problèmes type et des solutions" (1970, p.11). Ils ont ainsi une fonction normative et façonnent la vie scientifique en période de « science normale ». En période de crise, on verrait se multiplier les propositions concurrentes de théories avant que l'une ne vienne s'imposer et définir un nouveau paradigme directeur (exemple : passage du géocentrisme à l'héliocentrisme). Toutefois la polysémie du concept de paradigme est grande dans les écrits de Kuhn lui-même si l'on en croit l'analyse de Masterman (1970) qui a proposé une catégorisation de ces différentes significations en méta-paradigme, organisation sociale et art de faire. Dans son analyse rétrospective, Kuhn reconnaissait lui-même deux sens différents du terme dans ses écrits, et proposait de réserver le terme de paradigme au sens resserré de « solutions d'énigmes concrètes qui, employées comme modèles ou exemples, peuvent remplacer les règles explicites en tant que bases de solutions pour les énigmes qui subsistent dans la science normale. » (Kuhn, 1972, p. 207, cité par Willet, 1996), ou au sens d' « exemple commun » (Idid., p. 255) : « Bien que ce soit le sens de « paradigme » comme exemple canonique (« standard example ») qui m'ait originellement conduit au choix de ce terme, malheureusement, la plupart des lecteurs de *La Structure* ont manqué ce qui était sa fonction centrale, et usent du terme « paradigme » dans un sens proche de celui pour lequel je suggère désormais l'expression « matrice disciplinaire » (1990, p. 298, cité par Laugier, 2003, p. 499).

Dans la matrice disciplinaire, Kuhn distingue les lois scientifiques et leur formalisation, la conception du monde et les procédés heuristiques, les valeurs qui soudent le groupe des chercheurs, le modèle de résolution des problèmes (le paradigme servant d'exemple).

Remarquons également que c'est souvent la définition élargie qui est retenue par les auteurs du domaine de l'évaluation, soit l'ensemble des valeurs, données, méthodes et façons de voir le monde que partage une communauté de chercheurs. Il n'est pas étonnant que cette définition apparentée donc plutôt à la notion de matrice disciplinaire ait été retenue de manière large dans le champ des SHS, car qui pourrait dire que réellement une discipline des SHS (fut-elle une micro-communauté) se structure autour d'un paradigme au sens strict de Kuhn ? C'est sans doute ce qui explique la diversité des critères de démarcation retenus pour repérer les « paradigmes » qui structureraient les champs des recherches.

Ainsi, les communautés scientifiques se structurent autour de « paradigmes », entendus comme modèles de traitement des problèmes scientifiques, que l'on peut repérer à travers les manuels,

communications et pratiques scientifiques mais aussi plus largement comme « matrices disciplinaires » : ensemble des valeurs, données, méthodes et façons de voir le monde que partage une communauté de chercheurs et qui sont plus ou moins reconnus. La notion de « communautés épistémiques » (Haas, 1992 ; Adler et Haas, 1992) met l'accent sur la nature collective de la production de connaissances, mais aussi sur leur rapport au monde social, dans leurs tentatives d'influencer le politique. Ces communautés partagent un certain nombre de normes et de principes, de problèmes, de critères de validation des connaissances et des pratiques (1992). Plus précise que le terme discipline, cette expression semble mieux à même d'approcher les pratiques actuelles et concrètes des scientifiques. En la mettant en relation avec les notions connexes de communautés de pratiques (Wenger, 2000) ou de communautés de promesse (Brown, 2006), Meyer et Molyneux-Hodgson (2011) mettent l'accent sur leur caractère dynamique et transformatif. Elles se réalisent en même temps qu'elles produisent des connaissances. Elles façonnent, délimitent et articulent les identités de producteurs de connaissance actuels et futurs ainsi que les trajectoires individuelles et collectives.

Le problème épistémologique majeur mis en évidence par Kuhn (1962 trad. 1983) est celui de l'« incommensurabilité des paradigmes », c'est-à-dire de l'incomparabilité au moins partielle des théories rivales, étant donné que les chercheurs utilisent des langages incompatibles associés à des réalités différentes. Pouvoir accéder à la compréhension d'un autre paradigme que celui dans lequel on est immergé n'est possible qu'au prix d'efforts importants et de traductions toujours approximatives. Car « saisir, ou entrer dans un nouveau paradigme veut dire passer à une nouvelle façon de *voir* le monde. Il ne s'agit pas seulement d'*interpréter* différemment des données, mais de voir autre chose. Cette image perceptuelle, fondée sur la psychologie de la forme, conduit Kuhn de l'affirmation assez neutre que les savants appartenant à deux paradigmes différents voient les choses différemment, à l'affirmation plus effrayante qu'ils vivent dans des mondes différents. Tout cela passe, on le voit, par le concept de signification. » (Laugier, 2003, p.488).

Les apports de Kuhn sont fondamentaux en ce qu'ils montrent que la science est profondément située de manière historique et sociologique et qu'elle mêle des savoirs de différentes natures, dont des valeurs et des principes métaphysiques pouvant être antagonistes selon les communautés épistémiques. Bachelard avait déjà mis l'accent sur la mise en perspective historique de la science et sur la notion de rupture mais Kuhn a amené à envisager également l'incommensurabilité des paradigmes. Des discussions ont été soulevées quant au caractère universel d'un paradigme, que ce soit dans les sciences de la nature ou dans les sciences humaines et sociales. Nombre de développements contemporains font état de prolifération plutôt que de succession de régimes de vérité (Harding, 1991; Hollinger, 1994 ; Lather, 2006).

C'est également en ces termes de prolifération que nous pouvons aborder la recherche en évaluation à travers les différentes cartographies de générations de modèles.

1.2.3. Différentes cartographies de générations de modèles d'évaluation

Les tensions entre paradigmes conduisent à repositionner la recherche par rapport à leur pluralité et aux controverses que cela suscite dans leur utilisation comme dans leur classification. Non seulement il y a une multitude de modèles d'évaluation, rien que pour l'évaluation de programme, Stufflebeam en répertoriait déjà 21 en 2001, mais de plus, il semblerait bien qu'il y ait également autant de proposition de classements que d'auteurs. Dans l'ordre chronologique et sans exhaustivité, Madhaus, Scriven et Stufflebeam (1986) évoquent 6 âges ; selon Cardinet (1987) il s'agit de conceptions, externe, interne et négociée ; Shadish, Cook et Leviton (1991) considèrent trois générations ; Guba et Lincoln (1989) proposent une évaluation de 4^{ème} génération ; De Ketele (1993) identifie 11 paradigmes ; Fontan et Lachance (2005) reprennent les 4 générations de Guba et Lincoln et en proposent une 5^{ème}, Rodrigues (2006) identifie trois perspectives : objectiviste, subjectiviste, dialectique et critique ; Fitzpatrick, Sanders et Worthen (2011) repèrent 4 catégories d'approches ; Vial (2012) distingue trois grands modèles d'évaluation : mesure ; gestion et située ; Mertens (2015) explicite 4 paradigmes, différents des 4 paradigmes de Ketele (2016) (intégrant et prolongeant les 11 paradigmes identifiés en 2013) ; Younes (2018, 2020) identifie 3 générations et le défi d'une quatrième en cours.

Les auteurs proposant un méta regard sur les approches de l'évaluation s'accordent pour considérer qu'elles apparaissent progressivement dans le temps. On peut effectivement plus ou moins repérer une chronologie de ces approches, sans que les nouvelles effacent les anciennes. Elles se transforment, se combinent ou coexistent sans se remplacer comme dans le cas des sciences de la nature où l'héliocentrisme a remplacé le géocentrisme. Nous reprenons plus précisément les classifications proposées par Shadish, Cook et Leviton, par De Ketele, par Guba et Lincoln, et par Mertens.

- ***3 générations de modèles de l'évaluation des programmes selon Shadish, Cook et Leviton***

Shadish, Cook et Leviton (1991) ont distingué aux Etats-Unis, 3 générations de modèles de l'évaluation des programmes, des années 60 aux années 90.

- La première génération, dans les années 60, est représentée par Scriven et Campbell qui misent avant tout sur la rigueur scientifique en cherchant une connaissance causale valide des effets des programmes. Ainsi, Campbell (1969, 1984) privilégie l'approche expérimentale. Scriven, quant à lui, met l'accent sur une approche multidimensionnelle pour contrôler les biais avec le but

principal de produire un jugement de valeur : « Ce qui est mauvais est mauvais, ce qui est bon est bon et c'est le travail des évaluateurs de déterminer ce qu'il en est » (Scriven, 1986, p. 19, cité par Shadish, Cook et Leviton, 1991). Scriven a défini une logique de l'évaluation aujourd'hui largement reconnue consistant, après avoir choisi un objet à évaluer, à définir des critères d'évaluation, des niveaux d'excellence (absolus ou comparatifs), à déterminer la performance de ce qui est évalué et enfin, à la comparer à ces niveaux d'excellence (Scriven, 1980 p.18).

- Dans les années 70, une deuxième génération, (autour de Weiss, Wholey, et Stake), devant l'échec de la première génération à produire de la connaissance utile pour améliorer les programmes, a proposé de nouveaux modèles plus pragmatiques visant l'utilisation de la connaissance produite par l'évaluation pour améliorer les programmes. Weiss (1987, Weiss et Bucuvalas 1980, cités par Shadish, Cook et Leviton, 1991) a insisté sur l'importance de prendre en compte les réalités politiques et sociales et de produire des connaissances qui serviront à éclairer les décideurs politiques. Wholey (1983, cité par Shadish, Cook et Leviton, 1991) a visé davantage ceux qui dirigent les programmes au niveau local pour en améliorer le management. Stake, quant à lui, afin de travailler au niveau des acteurs, privilégie le rôle de celui qui facilite l'effort de changement par les acteurs : « La plus grande contribution de l'évaluateur est d'aider les gens à découvrir les idées, réponses et solutions dans leur propre esprit » (Stake, 1976, p. 36, cité par Shadish, Cook et Leviton, 1991).

- Shadish, Cook et Leviton (1991) identifient une troisième génération, à partir des années 80, représentée par Cronbach et Rossi, cherchant à concilier la recherche d'une connaissance descriptive concernant l'usage et la programmation sociale avec les exigences d'une connaissance valide.

- ***Vers un cadre intégrateur par De Ketele des différents paradigmes de l'évaluation : du modèle objectiviste au modèle répondant***

Les écrits de De Ketele sont significatifs en ce qu'ils donnent à voir un cheminement de sa pensée quant à la modélisation de l'évaluation qui devient plus englobante au fil du temps. En 1993, il distingue 11 paradigmes comme « cadres implicites qui définissent des problèmes, des méthodes et des solutions légitimes pour une communauté scientifique » (1993, p. 59). A l'exception du premier, ces paradigmes sont tous associés à des auteurs marquants⁴ : Le *paradigme de l'intuition pragmatique* est celui de l'évaluateur de terrain, pour qui l'évaluation est un acte syncrétique, peu ou pas formalisé et fortement dépendant de la personne et de la situation. Le *paradigme docimologique* où l'évaluation est un problème de fiabilité de la mesure, est à la base des

⁴ Nous renvoyons aux textes de De Ketele pour plus de précisions.

développements de la psychométrie ou de l'éduométrie. Le *paradigme sociologique* est celui qui nous interroge le plus. Fondé autour des « concepts » du déterminisme social et de l'équité, il nous semble renvoyer à une approche sociologique particulière à savoir le paradigme critique bourdieusien, alors que d'autres approches sociologiques de l'évaluation existent. Il s'agit ensuite du *paradigme de l'évaluation fondée sur les objectifs*, dont les limites behavioristes ont conduit à la formulation du *paradigme de l'évaluation formative dans un enseignement différencié*, où la notion de régulation est centrale. Il distingue ensuite le *paradigme fondé sur la décision* et la régulation au niveau du système dont le modèle CIPP (Context, Input, Process, Product) de Stufflebeam et Shinkfield (1985) est l'emblème mondialement connu qui, élaboré dans le cadre de l'évaluation de programme a mis l'accent sur la notion d'évaluation au service de la décision certes mais également les questions systémiques avec l'évaluation non seulement des produits, mais également du contexte et des processus. Le *paradigme de l'intégration*, allait naître d'une réflexion critique des travaux précédents et donner lieu aux concepts de compétence et de pertinence, qui reconfigurent la réflexion éducative. Le *paradigme de l'évaluation centrée sur le consommateur* selon Scriven (1967, 1983) se caractérise par le primat donné au jugement de valeur mais avec une méthodologie rigoureuse pour définir les critères et standards d'évaluation et contrôler les biais. Le *paradigme de l'évaluation centrée sur le client* ou paradigme de l'évaluation répondante selon Stake (1976) est une tentative d'intégration des propositions antérieures formulées dans le champ de l'évaluation de programme. Il est aujourd'hui unanimement reconnu comme étant l'ancêtre des approches de l'évaluation participatives et collaboratives, avec la mise en évidence des référentiels multiples et les collaborations chercheurs /praticiens. Il distingue ensuite le *paradigme économique*, très utilisé dans le domaine de la « qualité » autour des notions, d'efficacité, de rentabilité et redevabilité,. Il termine sa présentation en indiquant l'émergence d'un *paradigme de l'évaluation comme processus systémique de régulation* avec l'idée d'une régulation introduite au cœur des dispositifs et des programmes et non seulement des apprentissages. Y sont également traités les apports des perspectives multiples des référentiels.

20 ans plus tard, De Ketele (2016) propose un cadre plus intégrateur de ces différents paradigmes qu'il organise dans deux champs différents (considérant comme Allal, qu'il convient de les distinguer) : celui de l'évaluation centrée sur le contrat didactique et celui de l'évaluation centrée sur le contrat institutionnel, et selon que l'évaluation est centrée sur les produits, classements, référentiels fermés ou sur les processus, l'accompagnement et les référentiels ouverts.

Cette catégorisation l'amène à finalement distinguer 4 paradigmes de l'évaluation : au niveau du contrat didactique, il distingue le *paradigme de l'intuition pragmatique* (évaluation des apprentissages, sommative, subjective, faite par l'enseignant avec une posture de contrôle) et le

paradigme de la régulation interactive (pour l'apprentissage, interprétative, partagée et avec une posture de reconnaissance). Au niveau du contrat institutionnel, il distingue le *paradigme du management du système éducatif* (top down, évaluations externes standardisées) et le *paradigme du leadership collectif partagé* (bottom up, internes, participatives et situées).

De Ketele souligne à juste titre le rôle fondateur de Stake dans la conceptualisation des approches participatives de l'évaluation. En effet, Stake est considéré comme le premier théoricien à avoir, à la fin des années 60, produit (avec Guba) une critique radicale du modèle objectiviste et proposé une alternative avec son modèle de l'évaluation répondante et ses apports méthodologiques concernant les études de cas (Greene et Abma, 2001). L'évaluation est répondante en ce qu'elle s'adapte aux réalités du terrain, mettant l'accent sur « observer et réagir » (Abma et Stake, 2001). « Une évaluation est répondante si elle s'oriente plus directement vers les activités du programme que vers ses intentions, répond aux demandes d'information du public et si, lors du rapport quant au succès et échec du programme, les différentes perspectives de valeur en présence sont prises en compte » (Stake, 1975, p.14). Il s'agit donc d'une évaluation flexible, qui reconnaît et rend compte des réalités et des jugements de valeur multiples, donne plus de valeur aux théories et connaissances locales avec les nuances et les divergences, privilégie les études de cas et les méthodes qualitatives. Cette évaluation fait avec la complexité et la diversité plutôt que de tenter de les réduire et de les simplifier. Greene et Abma considèrent que Stake a permis de reconceptualiser l'évaluation de programme : d'une centration sur les questions des décideurs, distants et préoccupés d'efficacité et de bénéfice moyen, avec des techniques d'analyse sophistiquées, à une centration sur le programme in situ et un engagement avec les praticiens quant à la qualité et le sens de leur action (2001, p.1).

- ***Une quatrième génération d'évaluation avec Guba et Lincoln***

Guba et Lincoln sont allés encore plus loin dans la théorisation en approfondissant les soubassements philosophiques et épistémologiques de leur modèle dans leurs deux ouvrages de référence non traduits en français à notre connaissance : *Naturalistic inquiry* (1985) et *Fourth generation evaluation* (1989). Ils ont été déterminants dans l'élaboration du paradigme constructiviste de l'évaluation. Ils ont élaboré de nouveaux critères pour juger de la qualité des recherches et des évaluations, afin de pouvoir questionner la valeur de vérité des résultats produits, leur généralisabilité, leur consistance et la neutralité du chercheur. Dans les paradigmes positivistes et postpositivistes, les critères de qualité définis pour répondre à ces questions fondamentales sont successivement ceux de validité interne, validité externe, fidélité et objectivité. Guba et Lincoln proposent comme critères alternatifs à ceux élaborés par les chercheurs quantitativistes, la crédibilité (les enquêtés jugent-ils les propositions du chercheur

comme étant crédibles ?), la transférabilité (possible dans une certaine mesure), la fiabilité (prenant en compte les changements liés au contexte) et la confirmation (dans l'enquête prolongée). Selon eux, le rôle principal de l'évaluation consiste à répondre aux demandes d'informations du public de manière à rendre compte des différentes perspectives de valeur des parties prenantes. « L'enquête naturaliste » consiste à étudier le programme in situ tel qu'il se déroule, sans le contraindre, le manipuler ou le contrôler. L'évaluateur apprend des évalués. Guba et Lincoln (1989), ne définissent pas les générations de l'évaluation de la même manière que Shadish et al (1991). S'ils s'accordent pour considérer que la première génération de l'évaluation s'inscrit dans le paradigme de la mesure, selon eux la deuxième génération s'appuie sur la description pour comprendre les raisons des écarts observés entre les objectifs et les résultats afin d'apporter des régulations dans le programme, alors que la troisième génération réhabilitait la valeur subjective de l'acte d'évaluer, le rôle de l'évaluateur consistant à se prononcer sur la valeur et les mérites de l'objet évalué à partir de critères d'efficacité préalablement identifiés. Si chaque génération a apporté un progrès dans la conception et les méthodes de l'évaluation, Guba et Lincoln considèrent que ces trois générations se situent dans un paradigme post-positiviste (1994) ou rationaliste (1982) établi pour les sciences de la nature et qu'ils jugent inadapté aux sciences humaines et sociales. Ils ont proposé de fonder la quatrième génération d'évaluation sur un nouveau paradigme naturaliste (1982) ou constructiviste (1994), qu'ils assimilent à une révolution scientifique au sens de Kuhn, la rupture reposant sur le passage d'un réalisme ontologique à un relativisme ontologique. En effet, pour Guba et Lincoln (1994), et nous partageons leur position sur ce point, la question des méthodes est secondaire à celle des paradigmes qu'ils définissent comme un système de croyances qui guident le chercheur non seulement quant à ses choix méthodologiques, mais plus fondamentalement quant à ses positionnements ontologiques et épistémologiques. Il s'agit d'une vision du monde qui définit la nature du monde (1994, p.105). Ils analysent 4 paradigmes au fondement des approches des SHS, qu'ils distinguent sur le plan de leurs ontologies, de leurs épistémologies et de leurs méthodologies. Leur catégorisation de 1994 a été complétée en 2005 avec le paradigme participatif/coopératif de Heron et Reason (1997) et élargie à l'axiologie comme croyance centrale au fondement des paradigmes au même titre que l'ontologie, l'épistémologie, et la méthodologie. Leur travail a été poursuivi par Mertens (2009, 2015) et Mertens et Wilson (2012)

- Le positivisme serait fondé sur une ontologie que Guba et Lincoln (2005) qualifient de « réalisme naïf », selon laquelle il existe une réalité tangible, caractérisée par des lois et des mécanismes que l'enquête peut mettre au jour. La connaissance de cette réalité prend la forme de lois de cause à effet, décontextualisées. La posture de recherche est

déterministe et réductionniste. L'épistémologie est dualiste et objectiviste, c'est à dire que le chercheur et l'objet de l'enquête sont considérés comme des entités séparées, le chercheur peut étudier l'objet de manière objective, il faut éliminer les biais. La méthodologie repose sur l'expérimentation aléatorisée pour vérifier les hypothèses. En termes d'axiologie, la connaissance propositionnelle sur le monde est une fin en soi, intrinsèquement valable.

- Le *postpositivisme* est fondé sur une ontologie qualifiée de « réalisme critique » (Cook et Campbell, 1979) selon laquelle il existe une réalité mais qui ne peut être connue que de manière imparfaite. Au niveau épistémologique, si le dualisme est abandonné, l'objectivité reste un idéal régulateur obtenu par la tradition fondée sur la critique de la communauté scientifique et la vérification de l'adaptation des résultats aux connaissances existantes. Au niveau méthodologique, l'accent est mis sur la triangulation des données et la théorie ancrée, les enquêtes réalisées sur le terrain collectant plus d'informations situationnelles et sollicitant les points de vue des acteurs, pour identifier les buts et les significations qu'ils attribuent à leurs actions. Dans ce cadre cependant, l'expérimentation aléatorisée (ou quasi expérimentale) reste privilégiée pour l'administration de la preuve, à savoir déterminer l'efficacité d'une méthode ou d'un programme. En termes d'axiologie, la connaissance propositionnelle sur le monde est une fin en soi, intrinsèquement valable. Aux Etats-Unis, les principes éthiques et lignes de conduite pour la protection des sujets humains de la recherche (bienfaisance, respect et justice), formulés dans le rapport Belmont (1979), se sont traduits en des procédures comme la confidentialité, le consentement éclairé et le refus de la contrainte (Mertens et Wilson, 2012).
- La *théorie critique* est fondée sur une ontologie « réaliste historique » structuraliste ; différents types de facteurs sociaux, politiques, culturels, ethniques et de genre se cristallisant dans des structures « réelles », une réalité virtuelle ou historique. L'épistémologie est transactionnelle et subjectiviste. Investigateur et investigués étant liés, leurs valeurs influencent l'enquête. Les résultats sont médiatisés par les valeurs. La méthodologie est dialogique et dialectique. La nature transactionnelle de l'enquête requiert un dialogue entre chercheurs et sujets de la recherche, un dialogue qui doit être dialectique pour transformer l'ignorance et les incompréhensions en une conscience plus informée. Il est question d'« intellectuels transformateurs » (Giroux, 1985 ; Burns, 1978). En termes d'axiologie, la connaissance propositionnelle transactionnelle est valable instrumentalement comme un moyen d'émancipation sociale qui comme fin en soi est intrinsèquement valable. Dans une brève histoire du concept d'émancipation, Broussal (2017) en rappelle les différentes acceptions selon les époques ; du « gouvernement de

- soi », du « projet culturel », de la déconstruction critique des mécanismes de domination jusqu'à la proposition politique des dispositifs de recherche intervention (Marcel, 2017).
- Le *constructivisme* est fondé sur une ontologie relativiste. Les réalités sont appréhendables sous la forme de constructions mentales, multiples et intangibles de nature sociale et expérientielle, locales et spécifiques, même si des éléments communs sont souvent partagés. Dépendantes dans leurs formes et dans leurs contenus des personnes et des groupes, elles ne sont pas vraies mais plus ou moins informées et sophistiquées. Les constructions sont altérables comme leurs réalités associées. L'épistémologie est, comme dans le cas de la théorie critique, transactionnelle et subjectiviste. L'investigateur et l'objet de l'investigation sont considérés comme inextricablement liés de telle sorte que les « résultats » sont littéralement créés par l'investigation. La distinction conventionnelle entre l'ontologie et l'épistémologie, disparaît, comme dans le cas du paradigme critique. La méthodologie est herméneutique et dialectique. La nature variable et personnelle (intra individuelle) des constructions sociales suggère que les constructions individuelles peuvent être obtenues et affinées seulement à travers les interactions entre et parmi les enquêteurs et les enquêtés. Ces constructions sont interprétées via des techniques herméneutiques et sont comparées via un échange dialectique. Le but final est de distiller une construction consensuelle plus informée. En termes d'axiologie, comme dans le cas de la théorie critique, la connaissance propositionnelle transactionnelle est valable instrumentalement comme un moyen d'émancipation sociale qui comme fin en soi est intrinsèquement valable.
 - Le *paradigme participatif* est fondé sur une ontologie participative, la réalité étant co-créée par l'esprit et le monde donné. L'épistémologie s'appuie sur la subjectivité critique dans la transaction participante avec le monde. L'épistémologie est étendue à la connaissance expérientielle, propositionnelle et pratique, les résultats sont co-crés. En termes de méthodologie, l'accent est mis sur la participation politique et la recherche action collaborative. La primauté de la pratique et l'utilisation du langage sont ancrés dans le contexte expérientiel partagé. En termes d'axiologie, la connaissance pratique à propos d'un épanouissement équilibré entre autonomie, coopération et hiérarchie est une fin en soi intrinsèquement valable.

- ***Deux paradigmes, pragmatique et transformatif, ajoutés par Mertens à la classification de Guba et Lincoln***

Mertens (2015) propose une reformulation assez proche de celle de Guba et Lincoln en ce qui concerne les paradigmes post positivistes (dans lequel elle englobe le positivisme) et constructiviste mais elle identifie deux paradigmes supplémentaires :

- Le *paradigme pragmatique* sous-tendant les travaux de la deuxième génération d'évaluations selon Shadish, Cook et Leviton (1991), met l'accent sur l'utilisation des résultats. Il s'ancre sur les travaux de James, Dewey, Mead et Bentley, qui rejetaient l'idée que la vérité puisse être découverte par des méthodes scientifiques. Dans les années 60, les neopragmatistes Kaplan, Rorty et West se distinguent des premiers pragmatistes en mettant l'accent sur le sens commun et la pensée pratique (Mertens et Wilson, 2012). Mais plusieurs auteurs attirent l'attention sur la confusion fréquente entre le pragmatisme qui s'inscrit dans un cadre philosophique et l'utilitarisme dont le registre est celui de l'action (Hall, 2013, Greene, 2007, Denzin, 2012). Selon Mertens, cette tension a conduit à préciser l'hypothèse axiologique en lien avec la théorie utilitariste de l'éthique selon laquelle la valeur de quelque chose est fonction de ses conséquences (Christian, 2005).
- Le *paradigme transformatif* met l'accent sur la justice sociale. Selon Mertens et Wilson (2012), il s'origine dans les travaux de Kant et Hegel, qui ont initié l'importance de la réflexion critique sur les relations de pouvoir, et se prolonge aujourd'hui dans une multitude de perspectives théoriques (néomarxistes, féministes, genre, indigéniste et postcoloniales selon Harding (1998), qui contribuent à éclaircir comment les questions de pouvoir et de justice et comment l'économie, la race, la classe, le genre, les idéologies, les discours, l'éducation, la religion et les autres formes d'institutions sociales et de dynamiques culturelles interagissent et construisent un système social. Mertens (2009, p.4) définit le paradigme transformatif comme « un parapluie métaphysique qui réunit différents fils philosophiques ». Il est applicable aux gens qui subissent discrimination et oppression, et à l'étude des structures de pouvoir qui perpétuent les inégalités sociales. Dans ce paradigme, l'évaluateur a un mandat explicite pour travailler à transformer le statu quo (Ponterotto, 2005). Mertens (2015, p. 81) rappelle le positionnement de l'association américaine d'évaluation qui en 2011 a publié une déclaration sur la compétence culturelle (*Statement on cultural competence*)⁵, qui appelle les évaluateurs à s'engager dans une démarche d'apprentissage permanente pour éviter d'être aveugles à

⁵ <https://www.eval.org/p/cm/ld/fid=92>

leurs propres présupposés qui diffèrent selon les backgrounds : « La compétence culturelle requiert une conscience de soi, une réflexion sur ses propres positions culturelles et une conscience de celles des autres, et la capacité d'interagir de manière authentique et respectueuse avec les autres » (Mertens, 2015, p. 81). Ainsi, comme le constructivisme, le paradigme transformatif reconnaît la nature multifacette de la réalité en fonction des positions de chacun. La responsabilité de l'évaluateur est de concevoir des dispositifs permettant l'expression et la visibilisation de ces différences donc leur dévoilement. La connaissance n'est ni absolue, ni relative, elle est créée dans un contexte de pouvoir et de privilège. Les évaluateurs doivent développer des relations respectueuses et collaboratives qui sont culturellement répondantes aux besoins des différents groupes de parties prenantes dans le but d'établir des conditions propices à révéler la connaissance à partir de différentes positions. Les tensions peuvent émerger de ces différences de pouvoir et le challenge est de travailler avec des questions sensibles relatives aux questions de discrimination et d'oppression. Les méthodes mixtes sont souvent utilisées dans le but d'établir une relation dialogique et une compréhension contextuelle profonde. Les méthodes qualitatives et quantitatives peuvent être utilisées ensemble parce qu'elles révèlent différents aspects des phénomènes à l'étude et correspondent aux différents besoins d'information des diverses catégories d'acteurs. La méthodologie de l'enquête évaluative intègre une conception « cyclique » selon laquelle des résultats intermédiaires peuvent être utilisés. Un suivi est nécessaire pour augmenter le potentiel des résultats de l'évaluation de programme à renforcer les droits humains. Ce paradigme transformatif supporte la crédibilité de l'évaluation en ce que les parties prenantes sont impliquées au long du processus et que leurs perspectives tendent à être reflétées de manière respectueuse.

1.3. Controverses autour des paradigmes en jeu

Ainsi, le concept de paradigme permet de mettre en perspective les différentes catégorisations ou cartographies, tout en restant lui-même polysémique entre une acception rationaliste resserrée et une approche culturelle beaucoup plus large. Toutefois, il n'en demeure pas moins que toutes les pratiques scientifiques sont soumises à des croyances, des mythes, des standards et des normes qui fondent le consensus dans un groupe de chercheurs et déterminent le choix des problèmes étudiés et des méthodes retenues pour trouver des solutions. De manière plus ou moins latente ou explicite les paradigmes mobilisés, dans les deux acceptions élargies ou resserrées, se révèlent déterminants dans une approche épistémologique de l'évaluation.

1.3.1. Lutte des paradigmes

Bien que selon Guba et Lincoln (1994), l'expression de « guerre des paradigmes » proposée par Gage (1989) soit exagérée, ils adhèrent cependant à la thèse de l'incommensurabilité des paradigmes. Plus précisément, les paradigmes positivistes et néopositivistes seraient compatibles mais incompatibles avec les paradigmes constructivistes, participatifs et critiques qui seraient commensurables les uns avec les autres. Étant donné l'incompatibilité des postulats fondamentaux quant à l'existence ou non d'une réalité, la possibilité ou l'impossibilité d'une science objective, ils ne sont pas conciliables, et seule l'élaboration d'un « méta paradigme » pourrait résoudre leurs antagonismes fondamentaux. Guba et Lincoln s'opposent en cela à la position pragmatique défendue à l'époque par Gage (1989). Selon cet auteur, les différents paradigmes ne sont pas incompatibles et peuvent non seulement coexister de manière harmonieuse et respectueuse mais également s'enrichir mutuellement voire être mobilisés dans un même programme de recherche. Dans cet article prospectif, Gage souhaitait l'avènement d'une arène harmonieuse et productive, au service du progrès mais il redoutait deux scénarios alternatifs. Selon le premier scénario, le paradigme « postpositiviste » disparaissait sous le feu des critiques portées par les tenants des paradigmes alternatifs (interprétativiste/critique). Selon un autre scénario, la guerre se poursuivait entre communautés de chercheurs retranchés dans des camps assiégés repoussant les agressions de leurs adversaires.

Il est intéressant de se pencher sur cet article 30 ans après et de se demander quel scénario est advenu. Le paradigme positiviste n'a pas disparu, bien au contraire, et les mots de Guba et Lincoln (1994, p. 106) affirmant que la recherche quantitative se voit attribuer une plus grande valeur scientifique que la recherche qualitative restent d'actualité.

Avec le mouvement international de la recherche en éducation « basée sur la preuve », on ne peut que constater le retour en force de l'approche néo-positiviste comme seule approche scientifique réellement valable, balayant d'un coup 50 ans de recherches en philosophie et épistémologie des sciences sociales. En France, dans la présentation du premier colloque du Conseil scientifique de l'Éducation nationale et la composition des intervenants, le ton est donné. Il s'agit « de promouvoir la notion d'éducation fondée sur des preuves » car « Nombreux sont les pays à adopter le concept d'éducation fondée sur des preuves (evidence-based education), c'est-à-dire le fait de choisir les politiques éducatives en s'appuyant sur des données scientifiques solides ». Les intervenants sont économistes, statisticiens, neuroscientifiques, psychologues cognitivistes et les thématiques sont les suivantes : «

- « Les études randomisées contrôlées, qui permettent de comparer avec rigueur les progrès de deux groupes d'enfants en fonction de la stratégie pédagogique employée.

- Les grandes études statistiques nationales et internationales, qui fournissent une image détaillée des performances de milliers d'enfants et les analysent par des méthodes apparentées à celles de l'épidémiologie et de la sociologie.
- L'utilisation conjointe des méthodes de la psychologie expérimentale et de l'imagerie cérébrale, afin de comprendre les mécanismes de l'apprentissage et de mesurer les progrès des enfants.
- Le rôle de logiciels tels que le Graphogame ou La course aux nombres, à la fois pour diagnostiquer les différentes sources de difficultés en lecture ou en mathématiques, et pour proposer des remédiations sous forme de jeux. »

On ne peut que constater la disparition des paradigmes alternatifs dans de telles orientations choisies. Ceux-ci n'ont pourtant pas disparu. Ils se sont au contraire multipliés, témoignant de l'intensité de la vie scientifique dans un « paysage deleuzien de milliers de minuscules paradigmes » (Lather, 2006, p. 43). On serait face à un scénario où le primat politique et institutionnel du néo-positivisme renforcé par la puissance numérique s'accompagne d'une multitude d'alternatives situées.

Mais sans une claire conscience des enjeux scientifiques au sens large, les débats risquent de se réduire à la question des méthodes globalement opposées en méthodes quantitatives versus méthodes qualitatives. Certaines revues refusent les recherches fondées sur l'emploi des méthodes mixtes, n'acceptant que les recherches quantitatives ou qualitatives. Cependant de plus en plus, les méthodes mixtes sont acceptées et même préconisées (Tashakkori et Teddlie, 2003 ; 2009) pour des recherches alliant visées explicatives et visées compréhensives. Il est important de préciser qu'en relation à l'incommensurabilité des paradigmes, ce n'est pas le fait d'utiliser les « méthodes mixtes » qui résout le problème des paradigmes en jeu (Lincoln et Guba, 1985 ; Shadish, 1998 ; Mertens et Wilson, 2012). Dès les années 80, Lincoln et Guba affirmaient que dans le cadre d'un paradigme naturaliste /constructiviste, il était tout à fait possible et même souhaitable dans certains cas d'utiliser des méthodes quantitatives, même s'ils considéraient que les méthodes qualitatives étaient globalement plus appropriées car mieux adaptées à l'utilisation de "l'humain comme instrument" (1985, p. 199). Dans le cas d'un paradigme néopositiviste, ce sont les méthodes qualitatives qui se retrouvent souvent reléguées à une position ancillaire, descriptive, exploratoire, étant considéré que la véritable évaluation réside dans l'étude quantitative qui montrerait de manière irréfutable la valeur d'une méthode pédagogique par exemple. C'est donc bien au niveau des axiomes ou postulats fondamentaux que les controverses se situent.

1.3.2. Quatre principales controverses selon Guba et Lincoln

4 principales controverses sont repérées par Guba et Lincoln (2005) mettant en jeu l'action, le contrôle de la recherche, les fondations de la vérité et de la connaissance, et la validité.

- Controverses relatives à l'action

Alors que dans le cadre d'un paradigme positiviste ou néopositiviste l'action est vue comme une forme de militantisme qui vient réduire l'objectivité de la recherche, au contraire, dans le cadre des paradigmes post-modernes (tels le constructivisme, le critique, le participatif ...), elle s'inscrit comme un élément à part entière de la recherche, qu'il s'agisse de transformation interne comme degré de conscience plus important ou externe comme changement social. Un changement important s'est opéré dans le cadre du paradigme constructiviste et dans le cadre du paradigme participatif avec le pas franchi entre l'objectif de compréhension et celui d'action sociale (Lincoln, 1998). Que ce mouvement soit lié à une stratégie pour faire face au constat récurrent de la non utilisation des résultats de la recherche ou à un engagement politique, dans les deux cas l'enjeu est celui d'un changement. En effet la connexion de la recherche, de l'analyse politique de l'évaluation et/ou de la déconstruction sociale avec l'action, est devenue une caractéristique du travail d'enquête dans les nouveaux paradigmes à la fois au niveau théorique et au niveau théorico pratique.

- Controverses relatives au contrôle de la recherche :

Qui l'initie ? qui la conduit ? Qui en détermine les questions saillantes, la collecte des données, les résultats et leur diffusion ?

C'est une question qui ne se pose pas dans les paradigmes positivistes et postpositivistes, car la délégation du contrôle de la recherche aux sujets de l'enquête est vue comme une menace de la rigueur scientifique.

Au contraire, dans les paradigmes post-modernes, ce contrôle de la recherche impliquant les différentes parties prenantes ne met pas en cause la rigueur scientifique tout en étant un moyen de promouvoir l'émancipation, la démocratie et la puissance d'agir de la communauté. Ce n'est pas considéré comme une question de militantisme : « un terme quelque peu trompeur utilisé généralement comme un code dans un méta narratif visant à attaquer la rigueur de l'enquête » (p.174).

Cet aspect du contrôle de la recherche est en effet envisagé chez les tenants des théories critiques comme un moyen, pour les différents participants à la recherche, de prendre le contrôle de leur futur.

Pour les constructivistes, les différentes parties prenantes de la recherche sont amenées à participer à la définition des questions de la recherche et à la clarification de leurs implications pour des résultats partagés largement.

Dans le cadre des paradigmes participatifs, le contrôle de l'action par les participants de la recherche est clairement identifié comme un des buts de l'enquête.

- **Controverses relatives aux fondations de la vérité et de la connaissance.**

La question de savoir si le monde a une existence réelle en dehors de l'expérience humaine du monde ne se pose pas pour les positivistes elle car il est considéré qu'il y a une réalité extérieure. Et pour être mieux approchée, cette réalité doit être appréhendée avec des méthodes qui se prémunissent de la contagion humaine.

Au contraire, dans le champ des paradigmes critiques, constructivistes et participatifs ce sont précisément la connaissance sociale subjective et intersubjective, la construction active et la coconstruction de cette connaissance par les agents humains qui sont en jeu. Les différentes explorations sociales, intellectuelles et théoriques ont permis de mettre en évidence les inscriptions des auteurs autant que des lecteurs dans des contextes historiques et sociaux qui problématisent la vérité comme partielle, l'identité comme fluide, le langage comme système de référents non clair, les méthodes et critères comme potentiellement coercitifs. Réaliser la richesse des mondes mentaux, sociaux, linguistiques et psychologiques que les individus et groupes sociaux créent, recréent et co-crément constamment permet une forme de liberté plus grande dans l'enquête.

Il y a cependant des différences et des controverses entre les paradigmes. Pour les constructivistes, la vérité n'existe pas indépendamment des accords, négociations, dialogues, argumentations, discours qui s'élaborent entre les personnes et les groupes. Ce « concept de validité communicationnelle et pragmatique » (Rorty, 1979) n'est ni établi ni invariable. Il est construit par le moyen d'une communauté narrative elle-même sujette à des conditions temporelles et historiques qui permettent la création de la communauté. Il en va de même en ce qui concerne les débats sur l'évaluation et son appropriation ou son détournement par les acteurs concernés compte tenu des tensions, des décalages et des vérités « locales ».

- **Controverses relatives à la définition de la validité**

Les controverses au sujet de la définition de la validité se voient considérablement étendues. Il est important de considérer que la validité ne se confond pas avec l'objectivité. On a déjà pu discuter les fortes limites auxquelles la notion d'objectivité se heurte dans le champ des SHS. Selon Lincoln et Guba (2000, p.181), « l'objectivité est une chimère, une créature mythologique qui n'a jamais

existé, sauf dans l'imagination de ceux qui croient que la connaissance peut être distincte du connaisseur. »

Avec la validité, il s'agit de la question de l'authenticité des résultats. Présentent-ils une isomorphie suffisante avec les réalités dont ils sont censés rendre compte ? Sont-ils suffisamment dignes de confiance pour informer l'action ? L'histoire des sciences suggère en tout cas qu'aucune méthode ne délivrera une vérité ultime.

La question de l'interprétation est aussi très importante. Alors que les positivistes mettent l'accent sur la rigueur méthodologique, les post-modernes privilégient le consentement d'une communauté et une forme de rigueur (définie comme raisonnement défendable, plausible) dans la mise en avant d'une interprétation particulière et la conception d'une enquête interprétative (voir les critères définis). Il est rappelé qu'avec Schwandt (1996) et d'autres philosophes pragmatistes contemporains, l'enquête sociale est resituée, au-delà de la « critériologie » (p. 59), dans un cadre qui la transforme en une forme de philosophie pratique caractérisée par « des considérations esthétiques, de prudence et morales autant que des considérations scientifiques plus conventionnelles » (p. 68). Avec cette forme de « philosophie pratique », de questionnement approfondi sur comment se comporter dans le monde ainsi que sur les potentiels et les limites de la connaissance humaine, des critères totalement différents peuvent être définis pour juger de l'enquête sociale : apporter de l'information nouvelle plutôt que de chercher à confirmer des connaissances préétablies, accroître l'intelligence critique de tous et de chacun, juger l'enquêteur sur le potentiel de formation du jugement humain lié aux rapports qu'il produit.

La question n'est alors pas tant celle des critères que celle de la nature transformative de l'enquête. Transformation et critères sont ancrés dans des efforts dialogiques, eux-mêmes des formes de « discours moraux ». Lincoln et Guba explicitent trois facettes du concept de validité : authenticité, résistance/transgression poststructurale et relation éthique.

- *la validité comme authenticité* se décline en plusieurs critères : la justice/équité signifiant que toutes les perspectives, toutes les valeurs, toutes les voix doivent apparaître, y compris les minoritaires ; l'authenticité éducative et ontologique impliquant d'élever le niveau de conscience des participants à la recherche et autour d'eux ; l'authenticité catalytique et tactique déterminant la capacité d'un changement social positif et des formes d'action émancipatrices.

- *la validité comme résistance/transgression post structurale* est inspirée notamment par les travaux de Richardson (1997) qui appelle à une forme différente de science sociale amenant à repenser complètement la relation à son travail et à soi. Dans le but de « voir à quoi la transgression ressemble, il est nécessaire de trouver des méthodes qui nous permettent de découvrir les présupposés cachés et les répressions et dénis dans l'enquête de formes de vie. Re-

voir et re-raconter la sociologie sont inséparables » (p. 167). Richardson propose la notion de « validité cristalline » en lieu et place de la triangulation : "I propose that the central imaginary for "validity" for postmodernist texts is not the triangle : a rigid, fixed, two-dimensional object. Rather the central imaginary is the crystal, which combines symmetry and substance with an infinite variety of shapes, substances, transmutations, multidimensionalities, and angles of approach. Crystals, grow, change, alter, but are not amorphous. Crystals are prisms that reflect externalities and refract within themselves, creating different colors, patterns, arrays, casting off in different directions. What we see depends upon our angle of reponse. Not triangulation, crystallization. In postmodernist mixed-genre texts, we have moved from plane geometry to light theory, where light can be both waves and particles. Crystallization, without losing structure, deconstructs the traditional idea of ~validity" (we feel how there is no single truth, we see how texts validate themselves); and crystallization provides us with a deepened, complex, thoroughly partial understanding of the topic. Paradoxically, we know more and doubt what we know". (Richardson, 1997, p. 92). Cette proposition métaphorique fait écho à la validation catalytique de Lather (1986, 1993), une forme de validité qui refuse la complétude et favorise l'hétérogénéité, « une validité derridienne, rhizomique, avec des relais, des circuits, des ouvertures multiples » (1993, p.680), mais aussi une validité profondément située et partielle, associant éthique et épistémologie via des pratiques d'engagement et de self réflexivité (p.686).

- Dans cette perspective, Lincoln (1995) a *approfondi la validité comme relation éthique* et défini 7 standards : points de vue ; communautés discursives spécifiques et sites de recherche comme arbitres de la qualité ; plurivocité ; subjectivité critique (ou self reflexivity) ; réciprocité (plutôt que hiérarchie) ; sacralité (ou regard approfondi sur comment la science peut et doit contribuer au développement humain ; partage.

Ainsi les perspectives développées ouvrent des champs de réflexion problématiques ayant trait au problème des voix, de la réflexivité et des textes.

Au fur et à mesure que les chercheurs devinrent conscients du caractère complexe des réalités dont ils prétendaient rendre compte, ils commencèrent à introduire dans les textes les caractéristiques du langage oral des participants aux enquêtes et des chercheurs, mais aussi d'autres médias et d'autres formes (exemple poésie).

En ce qui concerne la réflexion critique sur le soi comme chercheur, il s'agit d'une prise en considération du soi comme non seulement chercheur mais aussi participant, enseignant et étudiant : celui qui apprend à connaître le soi dans le processus de la recherche elle-même. La réflexivité nous force à ne pas seulement traiter les questions de recherche en tant que telles mais également à considérer le soi lui-même et les identités multiples que représentent le soi

fluide dans le terrain de la recherche (Alcoff et Porter, 1993). Ainsi selon Reinharz (1997, p.3), non seulement on apporte le soi dans la recherche mais également on le crée. De même, l'écriture n'est pas une simple transcription de la réalité mais également un processus de découverte du sujet, du problème et de soi (Richardson, 1994).

Tout ceci enrichissant et complexifiant fortement les possibles, y compris au niveau des représentations textuelles où sont recherchés des formats qui échappent à la logique simplificatrice de la méthode scientifique, des « textes embrouillés » (Markus et Fischer, 1986) qui cassent la démarcation binaire entre la science et la littérature, pour broser la contradiction et la vérité de l'expérience humaine.

Après avoir espéré un temps que les chercheurs résoudre leurs différences et trouveraient un discours commun, Guba et Lincoln (2005) estiment que cette perspective est non seulement très improbable mais également non souhaitable. Ils considèrent que dans le moment postmoderne, poststructuraliste, il n'y a pas de vérité unique mais des vérités partielles : les glissements de sens entre signifié et signifiant créent des représentations qui ne sont que des fenêtres sur les individus, événements et places. De plus, le fait que les identités soient fluides plutôt que fixes, conduit à l'idée qu'il n'y a pas de paradigme unique qui peut être partagé en des termes identiques. Nous sommes plutôt au seuil d'une histoire marquée par la multivocité, les significations contestées, les controverses paradigmatiques et de nouvelles formes textuelles. Selon eux, il s'agirait aussi d'une émancipation de la « coercition de La Vérité » au sens d'Arendt, en tant qu'unique vision du monde. Ils font l'hypothèse de l'entrée dans un âge de plus grande spiritualité dans la recherche, ciblée sur des valeurs d'écologie, de respect des formes de vie non occidentales requérant une forte réflexivité en ce qui concerne l'attention au formatage occidental des enquêtes et à des enquêtes ciblant l'épanouissement humain (Heron et Rason, 1997). Cette quête pourrait aller jusqu'à réintégrer le sacré avec le profane dans les façons de promouvoir la liberté et l'autodétermination (Lincoln et Guba, 2000, p.186).

1.3.3. La prolifération des paradigmes avec Lather

En écho avec Harraway (1997, 2016), Lather (2006) insiste sur les multiplicités et invite à penser avec les tensions créées par la prolifération des paradigmes dans une période de résurgence positiviste et d'imposition politique du design expérimental comme étalon-or des méthodes de recherche. Considérant que le caractère linéaire des modèles structuraux vient « réduire et discipliner la profusion sauvage des choses existantes, elle propose un « mapping », que nous traduisons par cartographie. Dans une telle représentation, les catégories dualistes ne sont pas considérées sous l'angle de ruptures pures et simples mais comme des oppositions instables qui se déplacent et s'effondrent à la fois dans et entre les catégories. Ce qui conduit à tenter de

caractériser l'hybridation des positionnements contemporains et de lutter contre la rigidification des structures conceptuelles. Il ne s'agit pas de présenter une histoire linéaire du développement scientifique mais au contraire de maintenir délibérément les incompatibilités inévitables. Il s'agit de capturer le jeu des connaissances à la fois dominantes et émergentes en rivalité de légitimité « afin d'ouvrir une histoire de ce qui contient la pensée et comment la pensée est à la fois façonnée et dépassée par ce confinement ». Toujours selon Lather (2006, p. 43), « La cartographie des paradigmes peut nous aider à reconnaître à la fois notre nostalgie et notre désir de la circonspection vis à vis d'un foyer ontologique et épistémologique. La tâche est de savoir comment représenter le devenir de l'histoire contre les limites de nos cadres conceptuels qui sont encore beaucoup sur ce que nous avons déjà cessé d'être ».

1.3.4. L'hypothèse d'un paradigme synergique en évaluation

Au terme de cette analyse sur le positionnement épistémologique de l'évaluation dans les sciences de l'éducation et de la formation, l'accent peut être mis sur un des fils conducteurs inventoriés. Les éléments théoriques, critiques et opératoires sont mobilisés en leurs écarts et antagonismes dans les relations entre production de connaissances, mises en partage et transformations, dans un paysage de prolifération et de lutte des paradigmes.

A la suite du travail conduit sur les différentes classifications proposées par les chercheurs, leurs empiètements et leurs contradictions, j'en suis venue à formuler une nouvelle cartographie des différentes générations d'évaluation qui à la fois reprend les précédentes et s'en distingue, insistant sur des dominantes qui les caractérisent et amènent à progressivement enrichir les approches de l'évaluation (Younès, 2018, a et b, 2020). Nous reprenons l'hypothèse de Guba et Lincoln de 4 générations d'évaluation mais en les reformulant comme suit.

Le chantier de la première génération d'évaluations, qui a largement porté les efforts sur l'élaboration d'outils de mesure, d'explicitation de référentiels et d'indicateurs de performance individuelle et les tests standardisés se renforce en liaison avec les outils numériques et avec des méthodologies consolidées. Une deuxième génération prenant la mesure des limites du paradigme de la mesure a souligné l'importance d'associer mesure et régulation, quantitatif et qualitatif, standardisation et contextualisation, afin de disposer d'éléments objectivables situés. Une troisième génération a mis l'accent sur l'importance de la construction du sens individuel et partagé. Une quatrième génération d'évaluation écologique semble émerger, qui privilégie également les dimensions participatives et transformatives, en insistant sur un élargissement des dispositifs de recherche-action orientés par des valeurs de justice sociale et de prise en compte des différentes perspectives des parties prenantes, de telle sorte que le milieu d'évaluation soit considéré dans son ensemble.

La réflexion épistémologique sur les paradigmes coexistant à travers différentes générations conduit, dans le champ de l'enseignement universitaire, à mettre en relation des univers pluriels de recherche et de pratique en évaluation. Afin de n'évacuer ni l'incertain, ni les limites de la connaissance, ni les paradoxes et de se prémunir des simplifications, des systèmes clos et d'une orientation dualiste, il est visé de mobiliser et interroger les différentes strates engageant à **penser, construire, expérimenter un axe de recherche ouvert**. Ceci conduit à cibler les mises en relation, les interactions et inter-engendremets qu'implique une évaluation écologique. J'en suis ainsi arrivée à prospecter plus particulièrement dans quelle mesure et comment dispositifs d'évaluation et transformations peuvent s'articuler en faisant l'hypothèse d'un paradigme synergique. Il est ainsi considéré dans les chapitres suivants comment les dispositifs d'évaluation, en tant que dispositifs situés, coopératifs et appropriés, sont à penser moins comme procédures que comme dynamiques de champs relationnels, capables potentiellement de créer des conditions propices à des déploiements créatifs entre parties prenantes.

Partie 2 – Tensions au sein des dynamiques de l'évaluation : le cas de l'EEE

Rappelons tout d'abord des éléments de contexte permettant de situer la pratique de l'EEE (Evaluation de l'enseignement par les étudiants) et nos travaux au regard de la littérature internationale sur le sujet.

L'EEE désigne les pratiques par lesquelles les étudiants prennent part au jugement porté sur leurs enseignements. Cette démarche est devenue une norme institutionnelle dans un nombre croissant d'universités à travers le monde. Elle vise à améliorer le fonctionnement et la qualité pédagogique soit en fournissant aux enseignants et/ou aux structures pédagogiques le regard des bénéficiaires (nourrissant ainsi leur posture réflexive sur les enseignements et formations qu'ils conçoivent), soit en dotant les commissions de sélection ou de promotion de données fiables pour ancrer leur décision en bénéficiant d'un appui « objectif ». Elle est destinée, d'une manière implicite ou explicite, et selon la vision de l'institution, à la valorisation de l'enseignement universitaire et son adaptation. Les politiques de l'enseignement supérieur tendent de plus en plus à la mettre en avant comme levier. Ce mouvement n'est pas récent mais il prend un nouvel essor quand on le considère à la lumière de l'évolution des politiques de l'enseignement supérieur en Europe qui mettent l'accent sur la qualité, la rationalisation de l'université, la promotion de l'enseignement, la réforme des carrières académiques et la place centrale des étudiants.

Malgré des pratiques qui tendent à se développer, elle est toujours sujette à de vives controverses quant à son adoption ou son rejet, et ce pour des raisons différentes. Depuis sa mise en place, l'EEE a fait l'objet d'un nombre impressionnant de recherches (souvent anglo-saxonnes) donnant lieu au développement d'une littérature riche et divergente. En 1995, Cashin avait déjà répertorié plus de 1500 articles sur ce sujet conduisant Benton et Cashin (2012) à estimer qu'il s'agit de la thématique la plus étudiée dans le champ de la pédagogie universitaire, un résultat qui peut encore paraître étonnant en France où les recherches sur la thématique se sont pourtant développées ces dix dernières années. Les auteurs s'intéressent tant aux aspects épistémologiques et méthodologiques (pertinence, objectivité, critères, indicateurs, méthodes...) qu'aux effets (attendus ou inattendus) produits, la dynamique de l'EEE se trouvant prise entre différentes tensions : au niveau de sa validité, entre validité de la mesure et validité de conséquence, au niveau de ses finalités, entre démarche formative et démarche de contrôle, au niveau de l'adhésion des acteurs entre rejet, adhésion ou indifférence, au niveau de ses effets, positifs ou négatifs.

2.1. Tensions paradoxales

Ce sont ces tensions paradoxales comme oppositions productives que nous explorons dans cette première partie.

2.1.1. Entre validité de la mesure et validité conséquentielle

La plus grande partie de la littérature anglosaxonne foisonnante consacrée à cette thématique de l'EEE est consacrée à l'étude de sa validité dans une approche psychométrique. Il n'est pas question ici de passer en revue l'ensemble de ces recherches, les nôtres ne s'inscrivant pas dans ce paradigme, mais de donner un aperçu des méthodologies déployées et de souligner ce que l'on peut en retenir. Il s'agira aussi d'interroger la conception de la validité qui les sous-tend. Au niveau méthodologique Detroz (2008) relève les types d'études suivantes :

- des études multitraits-multiméthodes établissant des corrélations entre résultats obtenus par l'EEE et d'autres mesures de la qualité de l'enseignement ou selon différentes méthodes de recueil de données ;
- des études multisections qui consistent à comparer les scores à l'examen et la satisfaction des étudiants dans un contexte où les cours sont identiques (mêmes contenus, mêmes matières, mêmes notes de cours, mêmes évaluations – publics similaires en tous points) mais donnés par des enseignants différents ;
- des études analysant les biais potentiels de l'EEE. Dans ce cadre, les biais sont décrits comme des facteurs n'influençant pas les apprentissages mais influençant pourtant le score à l'EEE ;
- des études de laboratoire qui consistent à recréer une situation artificielle d'enseignement en maîtrisant certaines variables pour en limiter les effets.

Les méta-analyses effectuées sur les recherches basées sur ces méthodes tendent à fournir un faisceau significatif de données, interprétées le plus souvent comme étant des preuves de la validité de l'EEE. Même s'il s'agit de rester prudent, aucune des méthodologies décrites n'échappant à certaines critiques, elles s'avèrent précieuses car elles remettent en cause nombre d'idées reçues. A titre d'exemple, dans toutes les études de Marsh (1987), et contrairement aux résultats attendus, les enseignants demandant plus de travail aux étudiants obtiennent de meilleurs résultats que les enseignants leur demandant moins de travail. Il apparaît aussi que les enseignants indulgents ne sont pas nécessairement mieux évalués, et que les enseignants les plus sévères ne sont pas nécessairement les plus mal évalués (Ricci, 2004). Par ailleurs, si les enseignants expressifs, chaleureux et communicatifs sont plus appréciés, la recherche montre aussi que ces caractéristiques sont de puissants leviers de motivation des étudiants (Kim, Jörg et Klassen, 2019). Enfin, les EEE ne sont pas distinctes mais plutôt corrélées avec les appréciations d'observateurs extérieurs, d'anciens étudiants et d'enseignants.

Il a dès lors pu être considéré que les EEE étaient fiables, essentiellement dépendantes du comportement du professeur et de sa capacité à améliorer les apprentissages, et relativement peu affectées par les biais pressentis (Marsh, 1987, 2007). S'il nous semble important dans une perspective historique de rappeler cette conclusion plutôt en faveur de la pratique de l'EEE, son caractère général pourrait sembler quelque peu abusif. D'une part en masquant les variabilités situationnelles observées et d'autre part en considérant la validité sous un angle restreint alors même que dans les années 1980, les travaux de Lather (1986) sur la validité catalytique et ceux de Messick (1989) ont amené à reconsidérer le concept de validité en l'élargissant.⁶ Inspirée de la conscientisation de Freire (1974), la validité catalytique proposée par Lather est liée au développement des capacités des différents participants au processus de recherche évaluative. L'une des implications de cette reconsidération du concept de validité repose sur la prise en compte nécessaire, lors de son analyse, des conséquences (attendues et inattendues) du dispositif d'évaluation. Dans le domaine de l'EEE, cela revient à se poser la question de ses effets. C'est dans cette orientation, qui s'est développé dans les 20 dernières années, que nous inscrivons nos travaux.

2.1.2. Entre démarche formative et outil de contrôle

Les visées et effets de l'EEE ont provoqué de nombreux débats critiques et contradictoires à travers le monde. Il existe près de sept décennies de recherches et de réflexions sur cette thématique : dès les années 60 aux Etats-Unis, beaucoup plus récemment en Europe.

Bernard (1992), Centra (1993), Murray (1984), Doyle (1983), McKeachie (1979), Crawford et Shutler (1999), Detroz (2008), Younès (2006, 2009), Bedin (2009)... s'entendent sur les deux objectifs formatif et administratif, les plus importants de cette démarche : l'évaluation formative mise en place à des fins d'amélioration de l'enseignement, et l'évaluation administrative, réalisée à des fins de gestion de la carrière enseignante (promotion, sanction, renouvellement de contrats). Ces deux visées, plus ou moins distinctes traduisent de forts enjeux de nature différente ainsi que le montre le bref historique du développement institutionnel de l'EEE des deux côtés de l'Atlantique (Younès, Rege Colet, Detroz et Sylvestre, 2013). Il met en lumière les fluctuations des orientations politiques et les sérieuses difficultés pour en faire un instrument efficace au service de l'amélioration de l'enseignement universitaire, étant donné des usages non seulement explicites et annoncés officiellement, mais également implicites et détournés par les responsables locaux par exemple.

⁶ conduisant à une redéfinition reportée et actée dans les « standards for educational and psychological testing » (AERA, APA et NCME, 1999)

- **En Amérique du nord** : Si l'EEE est relativement récente en France, et irrégulièrement distribuée, elle est couramment utilisée en Amérique du Nord et plus généralement dans les pays anglophones depuis plusieurs décennies. Doyle (1983) en retrouve les premières traces au début du 20ème siècle quand, en 1926, les étudiants de l'Université de Harvard ont élaboré ce qui est considéré comme le premier questionnaire d'évaluation de l'enseignement par les étudiants. Mais l'utilisation de questionnaires a commencé à se répandre dans les années 1960 quand les étudiants ont réussi à obtenir une plus grande représentation dans les instances de gestion universitaire. L'origine de l'EEE est donc contemporaine d'une position plus active et engagée des étudiants par rapport à la vie académique. D'après Bernard (2011), ce mouvement s'est poursuivi dans les années 1970 avec une appropriation de la démarche par les enseignants dans une perspective d'évaluation formative. Elle note que, durant cette période et sous l'influence de professeurs innovants, le but de l'amélioration de l'enseignement s'est consolidé dans les universités nord-américaines, renforcé par la mise en place de centres de pédagogie universitaire offrant des services aux professeurs, en particulier des formations pour optimiser leur enseignement. Depuis, cette forme d'évaluation s'est fortement développée dans les universités américaines mais aussi canadiennes et québécoises : 29 % des collèges américains la pratiquaient en 1973, 68 % en 1983, 86 % en 1993 (Seldin, 1993) et, de nos jours, 90 à 100 % des collèges et universités américains (Trout, 2000).

Ce développement s'accompagne d'évolutions dans les politiques évaluatives. Dans les années 1970, les démarches institutionnelles d'EEE avaient principalement pour but de permettre aux enseignants d'améliorer leur enseignement. La visée formative de l'évaluation était ainsi priorisée. Or, dans les années 1980, avec les coupes sérieuses dans les budgets de l'enseignement supérieur, l'exigence de rendre des comptes en termes d'utilisation des ressources et d'efficacité s'est renforcée. La visée administrative ou de contrôle de l'évaluation s'est alors progressivement imposée. Les EEE ont ainsi été davantage utilisées dans les décisions prises pour le recrutement, la titularisation et la promotion des enseignants et l'attribution des cours que pour améliorer la stratégie pédagogique (Bernard, 2011).

En réalité, ces deux utilisations de contrôle et d'optimisation sont généralement confondues. Cashin (1996) a montré qu'il existe dans l'enseignement supérieur nord-américain un écart entre le discours affiché de régulation formative (avec comme but principal d'aider les enseignants à améliorer leurs performances) et les utilisations des résultats comme contrôle. Son étude indique que le principal effet est presque toujours de prendre des décisions vis-à-vis du personnel académique (embauche, promotion, augmentations). De même, Forest (2009) précise que, dans les universités québécoises, les EEE sont utilisées à des fins administratives, et ce alors que les politiques des universités ne sont pas toujours explicites à ce sujet.

- **En Europe** : L'Europe s'est intéressée à la démarche d'EEE plus tardivement. Dans le contexte français, elle n'a été mise à l'ordre du jour des réformes universitaires que dans les années 1990. Les lois Faure et Savary de 1984 qui ont introduit plusieurs réformes importantes dans l'enseignement supérieur français ne proposent aucune forme d'évaluation de l'enseignement à l'université. Les années 1990 marquent une rupture progressive de cette tradition et une injonction ministérielle de plus en plus soutenue, l'arrêté Bayrou en 1997 institutionnalisant l'obligation de cette évaluation de l'enseignement (Younès, 2006, 2007). C'est également à partir de la fin des années 90 et sous l'impulsion européenne, qu'elle se généralise dans les universités en Irlande (Moore et Kuol, 2005), Belgique francophone (Detroz, 2010), Allemagne (Schmidt et Loßnitzer, 2010), Suisse francophone (Rege Colet, 2009), Autriche et Suisse germanophone (Schnoz-Schmied, 2016). Il est admis par tous ces auteurs que le processus dit de Bologne car marqué par la déclaration de Bologne du 19 juin 1999, engageant la construction de l'Espace Européen de l'Enseignement Supérieur, doit être considéré comme un tournant. Cette déclaration définit, parmi les lignes d'actions, le développement d'instruments communs permettant une meilleure évaluation de la qualité de l'enseignement et notamment des agences nationales d'évaluation de la qualité de l'enseignement supérieur (AERES puis HCERES en France, OAQ en Suisse, AEQES en communautés française de Belgique...), produisant elles-mêmes des référentiels et des procédures. Cette orientation se confirme dans les différentes références publiées depuis près de deux décennies. Par exemple, la quatrième référence de l'ENQA (the European Association for Quality Assurance in Higher Education) exige de chaque établissement qu'il s'assure de la compétence et de la qualité de son corps enseignant. Les étudiants font partie du processus d'évaluation, en tant qu'experts dans les comités d'évaluation et dans les dispositifs d'EEE mis en place en interne dans les universités.

Mais que ce soit dans les textes européens qui conduisent à l'articulation de systèmes internes et externes de garantie de la qualité, ou que ce soit dans l'essor des procédures d'accréditation, on voit se dessiner des politiques d'évaluation plus orientées vers « rendre des comptes » selon une vision de l'évaluation-mesure, que vers « améliorer la pédagogie », même si le discours général associe ces deux visées. Dans sa recension des modèles de l'assurance-qualité, Van Damme (2003) identifie 4 approches, chacune fondée sur un type de mesure différent (excellence, cohérence des moyens et objectifs, présence de critères minimaux), ce qui souligne dans la diversité des conceptions de la qualité de l'enseignement la prédominance de la pensée de l'évaluation comme mesure. Cette orientation est manifeste à travers la production d'indicateurs de performance, souvent utilisés dans une perspective comparative. Les universités entrent dès lors dans un espace de contrôle fortement piloté par des organisations externes, espace propice à la standardisation des programmes et des méthodes. Deux constats apparaissent propres à

soutenir cette analyse en ce qui concerne l'EEE : le fait que soit privilégié le questionnaire unique et identique pour tous alors que d'autres démarches existent et d'autre part, que cette forme d'évaluation s'impose de manière hégémonique au détriment d'autres approches comme l'évaluation par les pairs, par des spécialistes de la pédagogie, l'autoévaluation... des méthodes à tout le moins complémentaires à l'EEE. L'évaluation des enseignements par les étudiants s'érige dès lors en norme, constituant plus « une fin en soi » qu'un moyen, un instrument au service de la régulation de l'enseignement.

La spécificité du système universitaire français repose sur le fait que les enseignants-chercheurs y sont fonctionnaires, que leur indépendance est garantie par la loi, et que leur évaluation est assurée par les pairs. L'EEE n'est pas un outil de gestion de la carrière des universitaires et la réflexion sur la prise en compte de l'enseignement dans la carrière reste actuellement à l'état de recommandations qui sont récurrentes dans la succession des rapports (Espéret, 2001 ; Fréville, 2001 ; Dejean, 2002 ; Bertrand, 2014), même si avec la réforme du statut des enseignants-chercheurs intervenue en 2009, le ministère dit vouloir prendre en compte l'investissement pédagogique dans la valorisation de la carrière des enseignants-chercheurs⁷. De ce fait, si l'EEE est rendue obligatoire, ses résultats sont « à destination des enseignants concernés » (arrêté Bayrou), justifiant le maintien de leur confidentialité. Plusieurs études de cas conduites en France et en Europe, au niveau de différentes universités, montrent les contradictions des responsables qui oscillent entre les différentes visées et se heurtent à la difficulté de faire de l'EEE un réel outil d'évolution pédagogique (Younès, Rege Colet, Detroz et Sylvestre, 2013). Cette tension, voire contradiction, se reflète dans la littérature entre ceux qui considèrent que l'utilisation à des fins de contrôle annule la démarche formative et ceux qui estiment que ces deux visées sont compatibles puisqu'elles concourent au même but, celui d'améliorer l'enseignement (Centra, 1993 ; Romainville, 2009).

Selon Marc Romainville, le constat généralisé de cette double face dans l'utilisation des EEE devrait conduire à mettre l'accent sur l'adhésion des acteurs, la pertinence des critères de qualité et la qualité du suivi qui sera donné à l'évaluation plutôt que sur cette distinction qu'il juge « confortable mais trompeuse ». Cette double face (Ardoino et Berger, 1989) semble constitutive de toute évaluation, ce qui amène à souligner l'importance de marquer les limites et les passages en jeu entre ces deux finalités. Le risque avéré du primat du contrôle implique de veiller à bien définir les domaines respectifs tout en accordant le plus grand soin aux conditions d'une

⁷ Article 7-1. Décret n° 2009-460 du 23 avril 2009 modifiant le décret n° 84-431 du 6 juin 1984 fixant les dispositions statutaires communes applicables aux enseignants-chercheurs et portant statut particulier du corps des professeurs des universités et du corps des maîtres de conférences et portant diverses dispositions relatives aux enseignants-chercheurs.

démarche formative. Il s'avère déterminant d'explicitier clairement la ou les visée(s) poursuivie(s), voire même de concevoir les dispositifs d'EEE en fonction de cette (ou ces) visée(s) préalablement définie(s) et acceptée(s). La distinction entre évaluation formative et évaluation de contrôle⁸ n'est pas qu'une affaire de détails, non plus qu'une distinction artificielle inhérente à un discours trop « pédagogisant », mais il s'agit d'une opposition déterminante car elle renvoie à des positionnements idéologiques, scientifiques, épistémologiques et méthodologiques fondamentalement différents en matière de conception de l'évaluation, qu'il s'agisse d'évaluation des apprentissages ou d'évaluation de l'enseignement. Lorsque l'on compare les démarches d'évaluation, on constate qu'elles se déclinent de manières différenciées en fonction des visées. Ardoino et Berger (1989) distinguaient le contrôle (conformité) de l'évaluation (appréciation). Berthiaume, Lanarès, Jacqmot, Winer et Rochat (2011) ont ainsi développé un outil permettant de positionner une institution en fonction de son approche de l'EEE plus ou moins orientée vers le contrôle ou le soutien au développement professionnel. Dans le cas d'une démarche de contrôle, les questions de validité, fidélité, équité et obligation devraient être particulièrement centrales. La démarche repose généralement sur un questionnaire préalablement défini sur la base des politiques de l'université en termes de formation, administré à des moments formellement établis et avec des feedbacks acheminés vers les autorités administratives. A l'inverse, une démarche formative vise une évaluation sensible et diagnostique, basée sur une approche optionnelle fondée sur un questionnaire qui a du sens pour l'enseignant (et donc au moins co-construit par lui-même) en fonction d'éléments qu'il est prêt ou non à réguler et des problèmes qu'il est prêt ou non à soulever. Ce questionnaire est administré selon un calendrier défini par l'enseignant en fonction de ses besoins et il est seul destinataire des résultats, afin d'éviter que les problèmes pour lesquels il cherche de l'information ne soient débattus sur la place publique. Forest (2009) propose une démarche d'évaluation intégrée dans laquelle les résultats des évaluations des enseignements sont agrégés pour apporter des informations aux responsables de formations ou, plus globalement, à l'institution (sans lever l'anonymat des enseignants), mais également pour améliorer la qualité de la formation.

2.1.3. Entre effets positifs, négatifs ou apparemment sans effet

Les recherches ont surtout étudié les effets de l'EEE, au niveau individuel des enseignants (Younès, 2007), qui peuvent dans une première catégorisation être distingués en effets positifs, négatifs ou apparemment sans effet.

⁸ Nous pourrions aussi parler de démarche certificative, soulignant le caractère validant de la démarche. Nous avons opté pour le terme générique de contrôle indiquant le caractère externe d'une démarche où l'évaluateur est extérieur à la situation d'enseignement ou d'apprentissage évaluée, avec ou sans visée certificative.

De nombreux effets positifs sur les pratiques d'enseignement sont rapportés

Les enseignants n'ont souvent qu'une connaissance approximative de la perception des étudiants à leur égard (les corrélations entre les évaluations par les étudiants et l'auto-évaluation des enseignants varient de 0.2 Centra, 1973 à 0,49 Marsh et Ware, 1982). Les enseignants pratiquant les EEE considèrent généralement qu'elles sont une source d'information utile pour améliorer leurs cours (Bernard, Postiaux et Salcin, 2000). Ce qui a été confirmé dès les années 80 par les recherches de Marsh et ses collaborateurs (Marsh, Fleiner, et Thomas, 1975 ; Overall et Marsh, 1979), par la méta-analyse de Cohen (1980) et a été vérifié de nombreuses fois par la suite en Suisse (Dumont, Rochat, Berthiaume et Lanarès, 2012) ; en Belgique (Salcin et al, 2012) ; en France (Perret, 2017).

Mais les EEE peuvent également avoir des effets négatifs. La croyance suivant laquelle un « feedback » sur la performance d'enseignement conduit nécessairement à l'améliorer est à remettre en question. Déjà en 1979, Mc Keachie montrait que de mauvaises évaluations répétées avaient pour effet de décourager l'enseignant et de le rendre anxieux, ce qui rendait l'évaluation peu susceptible d'avoir un impact positif sur la qualité de son enseignement. De même Machell (1989) et Boice (1992) ont décrit l'attitude dépitée, voire mélancolique, de jeunes enseignants qui, malgré leurs efforts, sont soumis à la critique des étudiants.

L'absence apparente d'effets semble également fréquente. Le manque d'effets des EEE sur l'enseignement est souvent relevé, voire déploré. Wright et O'Neil (1995) constataient que les autorités des universités nord-américaines, canadiennes, australiennes et britanniques classaient l'évaluation de l'enseignement en fin de session, presque en dernière position parmi 36 mesures proposées pour l'amélioration de l'enseignement (l'EEE en cours de session étant classée en 21^{ème} position). De fait, de nombreuses études montrent qu'il ne suffit pas de « faire » de l'évaluation pour qu'elle ait un impact positif (Centra, 1993, Seldin, 1993, Kember et al 2002). Il semble même y avoir un certain nombre d'éléments indiquant que la plupart de l'information recueillie n'est pas utilisée du tout (Desjardins et Bernard, 2002 ; Johnson, 1999 ; Beran et Rokosh, 2009).

Pour plusieurs auteurs, les effets négatifs ou apparemment nuls seraient provoqués par une méthodologie et une organisation institutionnelle inadaptées, notamment parce qu'elles n'apportent pas aux enseignants un soutien qui leur permettrait de tirer profit des évaluations (Zabaleta, 2007 ; Theall et Franklin, 2001). Il s'agit de méthodologies mais aussi de cultures et de valeurs. L'inscription de l'EEE dans une réelle politique de qualité de l'enseignement et donc de soutien au développement professionnel pédagogique de l'enseignant universitaire est en question. Dans les années 2000, une série d'écrits sont parus qui tentaient, à partir de la recherche, d'identifier les conditions institutionnelles favorables à une dynamique formatrice de l'EEE (Romainville et Coggi, 2009). Aujourd'hui, les recherches se poursuivent dans une approche

compréhensive et expérimentale. Il s'agit à la fois de comprendre les logiques d'acteurs dans les situations réelles et d'autre part d'expérimenter des dispositifs tentant d'opérationnaliser leur orientation formative.

Par ailleurs, d'autres recherches sont plus nuancées. Ainsi, Schmelkin et al. (1997), Detroz (2010), Perret (2016), constatent dans des contextes universitaires différents, respectivement aux États-Unis, en Belgique et en France, qu'environ 40% des enseignants déclarent utiliser l'EEE pour faire évoluer leur enseignement. De même dans l'étude de Beran et Rokosh (2009) qui ont pourtant souligné l'importance des critiques émises par les enseignants, 36 % ont émis des commentaires positifs au sujet de l'utilité des EEE pour des objectifs tant administratifs que formatifs. Il semblerait donc qu'à travers les contextes, une part quantitativement relativement stable, minoritaire mais néanmoins conséquente (plus du tiers) des enseignants considère que l'EEE est utile pour améliorer l'enseignement.

2.1.4. L'impact des représentations de l'évaluation

Dans la perspective d'en comprendre les effets sur la pratique enseignante, plusieurs auteurs ont étudié comment les enseignants percevaient cette démarche d'EEE, considérant que sa crédibilité était déterminante quant à son impact sur la transformation de l'enseignement (Costin, Greenough et Menges, 1971 ; DeNisi et Kluger, 2000). Wachtel (1998) et plus récemment, Moskal et ses collègues (2012), établissent un lien direct entre l'usage de l'EEE et les croyances et perceptions des enseignants.

L'importante littérature nord-américaine basée sur des témoignages donne le sentiment que les enseignants résistent à l'EEE parce qu'ils doutent de la qualité et de la légitimité des données (Nasser et Fresko, 2002 ; Schmelkin, Spencer, et Gellman, 1997 ; Wachtel, 1998). Trois types d'opinions sont couramment rencontrés : réduire les exigences en donnant moins de travail et plus de bonnes notes permet d'obtenir de meilleures appréciations de ses cours ; les EEE reflèteraient plus les capacités de communication de l'enseignant que son efficacité pédagogique ; les étudiants comprendront plus tard l'intérêt d'un cours qu'ils peuvent méjuger pendant leur formation. Même si ces opinions sont plutôt contredites par la recherche, comme nous l'avons vu précédemment, elles sont parfois tenaces chez les enseignants, ce qui fait dire à Cohen (1990) que « les attitudes négatives au sujet de l'EEE sont spécialement résistantes au changement et qu'il semble que les enseignants et les administrateurs renforcent leur croyance en certains mythes relatifs à l'EEE par des éléments personnels et anecdotiques qui, pour eux, ont plus de poids que les preuves apportées par la recherche empirique » (p.126). Comme le souligne Detroz (2016), d'autres auteurs comme Aleamoni (1999) et Theall (2002) parlent également de « mythes » lorsqu'ils évoquent ces opinions des enseignants qui selon Marsh (1987) ne reposent

généralement sur aucune base empirique, mais sont devenues des représentations collectivement partagées par les enseignants. Cette affirmation selon laquelle ces opinions des enseignants ne reposent généralement sur aucune base empirique paraît quelque peu abusive et fondée sur le paradigme scientifique de la mesure, elle-même approchée de manière statistique. Les récurrences et tendances mises en évidence dans ces recherches n'invalident pas pour autant les « exceptions » relevées dans la littérature (par exemple par Cohen et Benson, 1988), montrant que ces tendances peuvent se trouver contredites dans des situations données.

D'autres recherches ont investigué les attitudes des enseignants envers l'évaluation de l'enseignement par les étudiants, considérant également que des attitudes négatives pourraient expliquer le manque d'intérêt de certains pour les feedbacks produits par ce type d'évaluation. Plusieurs recherches ont étudié les attitudes des enseignants sur les buts, les bénéfices, les moyens, les sources d'information et les modalités de l'évaluation des enseignements par les étudiants avec des enquêtes par questionnaires, au Québec (Bernard et al., 2000), en Israël (Nasser et Fresko, 2002), en France (Younès, 2007). Detroz et Verpoorten (2017) concluent que ces trois études, malgré leur dispersion géographique dans trois continents, montrent des résultats relativement analogues : les représentations des enseignants vis-à-vis de l'EEE ne sont pas univoques, mais, de manière globale, plutôt positives même si une partie des enseignants restent réfractaires à la démarche. Ces études vont dans le même sens que celle de Beran et Violato (2005) qui font état d'attitudes généralement positives ou neutres concernant l'évaluation de l'enseignement par les étudiants. On peut se demander aussi si les opinions négatives se justifient ou sont motivées de la même manière.

Il faut cependant souligner que les enseignants sont beaucoup plus partagés quand on les interroge sur les dispositifs institutionnels. Par exemple dans la recherche de Younès (2006, 2007) seule une légère majorité a trouvé que l'EEE pouvait contribuer à améliorer la formation par la réflexion critique qu'elle suscitait et par l'incitation à modifier la pratique pédagogique. Les trois quarts de ces enseignants ayant considéré avoir été stimulés par l'EEE pour faire évoluer leur enseignement déclarent avoir décidé de modifications dans le contenu de l'enseignement, dans son organisation, dans leur style pédagogique et dans les modalités d'évaluation des étudiants. Un peu plus de la moitié a noté également que l'évaluation pouvait contribuer à rendre plus clairs les objectifs et les pratiques pour soi-même et pour l'équipe pédagogique. Notons par contre que quelques-uns ont craint que l'EEE n'entraîne au contraire une baisse de la qualité en suscitant un effet pervers, celui de la démagogie et d'autres se sont dits agacés ou n'ayant pas compris les évaluations.

Avec une méthodologie reposant sur des questions ouvertes posées à 357 professeurs d'une université canadienne, Beran et Rokosh (2009) ont révélé l'ampleur des critiques des enseignants

qui expriment surtout des commentaires négatifs sur le dispositif d'EEE en particulier l'outil d'évaluation, les procédures administratives et l'usage des résultats, le tiers des répondants évoquant aussi des évaluations biaisées ou des étudiants non aptes à évaluer l'enseignement. Sans oublier les 11 % exprimant des effets négatifs sur eux-mêmes comme du découragement ou une baisse du niveau des cours. Beaucoup moins de commentaires positifs étaient exprimés. Quant à Spiller et Harris (2013), ils concluent à un écart important entre les perceptions plutôt positives de l'EEE et son utilisation effective par les enseignants-chercheurs puisque seulement 12% des répondants à leur enquête déclarent utiliser les résultats pour améliorer leur enseignement. Pour Burden (2008) au Japon ainsi que pour Beran et Rokosh (2009) au Canada ainsi que pour Stein et al. (2012), la plupart des enseignants adoptent une posture d'acceptation passive vis-à-vis de l'EEE, ce qui se traduit par des attitudes légèrement positives ou neutres, mais conduit rarement à une utilisation pro-active de ses résultats. Nous avons également relevé une participation limitée des enseignants et l'importance capitale des responsables intermédiaires dans la réussite du processus (Younès, 2007, 2009). Cette posture de résignation fait obstacle à l'utilisation des feedbacks issus de l'EEE dans une perspective de développement professionnel. Il y a donc d'une part un décalage entre des attitudes générales plus positives que les effets réels de l'EEE et d'autre part de fortes divergences entre les enseignants qui amènent à questionner à la fois les facteurs contextuels nationaux et locaux ainsi que les trajectoires personnelles et le positionnement des acteurs : décideurs, enseignants et étudiants divergent et les décalages et tensions doivent aussi travailler en rapport avec cette réalité. La subjectivité des acteurs signifie également porter un jugement à partir des interactions directes et indirectes, réelles et symboliques entre les partenaires. Le travail de recherche que nous avons conduit en France (Paivandi et Younès, 2017 ; 2019) montre les dimensions liées à l'imaginaire collectif et à l'interprétation subjective de chaque enseignant qui donnent sens au fait d'être évalué par les étudiants.

2.1.5. Entre émotion, cognition et identité

Dans le contexte d'évaluation des établissements d'enseignement supérieur et de la qualité de l'enseignement qu'ils délivrent, les aspects psychosociaux de l'évaluation de l'enseignement sont une voie de recherche prometteuse. Nous avons déjà évoqué les travaux pointant, dans certains cas, des conséquences émotionnelles négatives de l'EEE, en termes d'anxiété voire de découragement (McKeachie, 1979 ; Machell, 1989 ; Boice, 1992 ; Johnson, 2000 ; Yao et Grady, 2005). Soulignons au passage que cette entrée par les émotions négatives nous semble liée à un rapport culturel aux émotions qui amène à plutôt les considérer sous l'angle des perturbations qu'elles induisent alors mêmes que le volume impressionnant de recherches consacrées à

l'émotion montre le rôle crucial des émotions dans le traitement de l'information qui nous provient du monde. Mais il y en a relativement peu consacrées à l'étude spécifique des relations entre les EEE et les émotions. Celles qui le sont procèdent par l'interrogation des enseignants quant à leur réception des EEE, en leur demandant de se centrer sur une expérience particulière et de se positionner quant aux émotions ressenties. Arthur (2009) et Barras (2017) concluent de leurs enquêtes que l'EEE induit des réactions émotionnelles chez toutes les personnes interrogées, plus souvent à valence positive que négative, mais avec de fortes variations interindividuelles. Kogan, Schoenfeld-Tacher et Hellyer (2010) interrogent 75 enseignants dans leur université américaine à l'aide d'un questionnaire en ligne, en proposant 18 émotions à positionner entre fortement ressentie et pas du tout. Leurs résultats indiquent qu'il y a peu de différences selon le statut des enseignants. Par contre, le genre a un impact, les femmes ressentant plus d'émotions négatives que les hommes. Barras (2016) utilise le QGA (Geneva Emotion Research Group, 2002) fondé sur cette théorie et les travaux de Scherer (2001) dans une haute école suisse. Les enseignants répondant à son enquête font part d'émotions variées à la suite de la lecture des résultats de leur dernière EEE (entre 0 et 9 émotions différentes avec des valences importantes et variées). La valence émotionnelle est neutre pour seulement 19 d'entre eux, alors qu'elle est positive pour 72 et négative pour 59. La valence émotionnelle semble être en moyenne plus négative chez les enseignantes que chez les enseignants. Le statut institutionnel des enseignants ne modifie pas le nombre d'émotions déclarées ni leur valence émotionnelle (sauf pour le nombre d'émotions chez les chargés de cours qui est moindre que chez les professeurs et pour la valence émotionnelle chez les chargés de cours plus faible que chez les responsables académiques). L'expertise en enseignement n'indique pas de différence notable dans l'intensité et la durée de l'émotion déclarée. Par contre, les novices réduisent moins la durée de cette émotion que les enseignants initiés et experts. Ce qui amène l'auteur à faire l'hypothèse qu'avec l'expérience, les enseignants développeraient une plus grande capacité ou besoin de contrôle de leur réaction émotionnelle. Il évoque également un « mécanisme de blocage émotionnel dans l'épuisement professionnel après plusieurs années de pratique. Il souligne que les commentaires négatifs rapportés par les enseignants sont émotionnellement forts, exprimant de la frustration, de l'injustice, voire de la colère et il propose que cette dimension soit prise en compte dans l'accompagnement des enseignants. Cette orientation de recherche plus rare est à poursuivre pour mieux identifier les antécédents et les conséquences de ces diverses émotions ressenties. Le modèle théorique de « l'appraisal » selon lequel les émotions apparaîtraient à la suite d'une évaluation de certains critères pertinents pour l'individu dans une situation donnée (Lazarus, 1966, Grandjeand et Scherer, 2009) qui permet d'expliquer pourquoi un même événement induit différentes émotions chez différents individus, selon leurs buts, leur culture,

leurs croyances et leurs capacités, semble être une voie prometteuse à mobiliser. Étant donnée la fonction adaptative des émotions, la compréhension de ces processus permettrait de mieux comprendre les processus interprétatifs et conatifs consécutifs aux EEE. Ces phénomènes émotionnels ont également partie liée avec les facteurs motivationnels ainsi qu'avec les dynamiques identitaires et évaluatives, elles-mêmes inter reliées.

- ***Attribution causale***

Sur le plan des attributions causales, en croisant la causalité interne ou externe, à savoir la mesure dans laquelle l'enseignant s'attribue le résultat négatif des évaluations où l'attribue aux étudiants, et le sentiment de contrôle qu'il a sur les régulations possibles, Arthur (2009) identifie quatre types de réactions possibles à ce feedback (Detroz et Verpoorten, 2017) :

- « Shame » : L'enseignant reçoit un feedback négatif à propos d'un aspect de son enseignement dont il s'attribue la responsabilité, mais qu'il ne pense pas pouvoir influencer ou modifier.
- « Blame » : L'enseignant reçoit un feedback négatif à propos d'un aspect de son enseignement pour lequel il juge les étudiants responsables et vis-à-vis duquel il ne compte pas agir.
- « Reframe » : L'enseignant reçoit un feedback négatif à propos d'un aspect de son enseignement dont il s'attribue la responsabilité et qu'il pense pouvoir influencer ou modifier.
- « Tame » : L'enseignant reçoit un feedback négatif à propos d'un aspect de son enseignement pour lequel il juge les étudiants responsables et vis-à-vis duquel il compte néanmoins agir.

Cette proposition qui met l'accent sur les processus d'attribution causale (responsabilité) et sur le sentiment de contrôle (possibilité d'agir), médiatisant l'engagement ou non de l'enseignant dans la régulation de l'enseignement, est à référer aux théories motivationnelles à partir desquelles on peut lui donner un sens et l'approfondir, en particulier la théorie de Bandura sur le sentiment d'efficacité et l'agentivité et la théorie de l'auto-détermination de Deci et Ryan (2002).

- ***Dynamiques identitaires et évaluatives***

Bien que l'identité professionnelle de l'universitaire soit pensée de manière de plus en plus complexe (Trowler, Saunders et Bamber, 2012), la forte association entre les activités d'enseignement et de recherche ainsi que la socialisation disciplinaire demeurent un fondement important (Burgel, 2006, Musselin, 2008 ; Musselin et Becquet, 2004 ; Taylor, 2008). Le fait que le recrutement et la carrière soient quasi exclusivement fondés sur les activités de recherche manifeste bien une hiérarchie des valeurs dans l'institution qui participe à la construction

identitaire des enseignants-chercheurs. Pour autant, ce primat permet-il de conclure que « les enseignants-chercheurs sont physiciens avant d'être enseignants de physique » selon une opinion couramment répandue (Endrizzi, 2017) ? Nous partageons plutôt la position de De Hosson, Decamp, Morand, et Robert (2014) pour qui : « tout se passe comme si l'enseignant-chercheur puisait dans son identité de chercheur les ressources pour bien enseigner ». Ils trouvent dans leur enquête que les enseignants-chercheurs valorisent autant l'enseignement que la recherche. D'ailleurs l'enquête récente de Chan, Luk et Zeng (2014) conduite dans une université de recherche intensive de Hong Kong (et plus spécialement en sciences, faculté classée dans le top 5 des universités asiatiques) utilisant les EEE pour des fonctions administratives et formatives montre que les enseignants sont moins de la moitié à déclarer préférer la recherche à l'enseignement et ils ne sont pas plus nombreux à déclarer passer plus de temps sur la recherche que sur l'enseignement. Ils sont 83 % à déclarer faire évoluer chaque année leur enseignement grâce aux EEE, 64 % à avoir participé aux activités du centre de ressources en enseignement et 1/3 à faire de la recherche sur l'enseignement. Dans notre enquête en France également, l'enseignement apparaît comme une dimension identitaire constitutive importante des enseignant.es-chercheu.r.es (Paivandi et Younès, 2019).

Par ailleurs, les recherches sur le « feedback » de performance (que l'on peut rapprocher des résultats d'une EEE) ont mis en évidence des effets variables suivant les facteurs interagissant, dont certains renvoient aux sujets et d'autres aux caractéristiques de la situation. Alors que l'on considère généralement que le feedback sur les performances peut permettre de les améliorer en fournissant à l'individu des informations susceptibles de lui servir pour ce faire, plusieurs recherches ont au contraire mis à jour les effets négatifs dus à la situation de comparaison sociale elle-même. Etre en situation de comparaison défavorable aurait un impact délétère, réduisant par là même les performances ultérieures. Il a été montré que les sujets à faible estime de soi sont particulièrement affectés par tout type d'information critique. Ils ont même tendance à surévaluer la portée des feedbacks négatifs qui font écho à leur manque de confiance en eux, et à donner moins de crédit aux feedbacks positifs (Brown, 2010). En revanche, les sujets à forte estime de soi acceptent seulement l'information consistante, c'est-à-dire l'information positive à leurs yeux (Campbell, 1990). Cette différence de traitement des feedbacks a récemment également été mise en évidence dans des travaux en neurosciences. Van der Miesen (2015) montre que les zones du cerveau activées suite à la réception d'un feedback négatif ne sont pas les mêmes selon que les personnes ont une faible ou à l'inverse une forte estime de soi. Pour Josephs, Larrick, Steele et Nisbett (1992), les sujets à faible estime de soi disposeraient de peu de ressources auto-protectrices et se comporteraient de façon à se protéger de toute menace. Les

individus à forte estime de soi feraient plus facilement face à cette menace du fait de ressources auto-protectrices abondantes. D'une façon générale, la stabilité/instabilité de la composante affective du soi semble dépendre, pour une part non négligeable, de la certitude que les individus ont à propos d'eux-mêmes (Banaji et Prentice, 1994). Plus cette composante est puissante (estime de soi forte), plus elle est stable et insensible aux variations situationnelles. Moins elle est forte, plus elle est instable et donc sensible aux caractéristiques situationnelles. Il est vrai que ces résultats pourraient également s'appliquer aux feedbacks issus de l'évaluation de la recherche. Ceux-ci n'ont pourtant pas suscité autant de controverses que les EEE, même si l'on observe une contestation grandissante des modalités d'évaluation de la recherche⁹. En suivant ces propositions, Detroz (2017) propose de croiser l'estime de soi et la valence des feedbacks selon le schéma suivant :

- Les enseignants à forte estime d'eux-mêmes recevant un feedback positif l'interprèteraient comme une confirmation de leur qualité d'enseignant. Ils ne seraient pas enclins à se servir des éventuelles critiques de l'enseignement apparaissant dans les EEE pour le faire évoluer.
- Les enseignants à forte estime d'eux-mêmes recevant un feedback négatif auraient tendance à le rejeter et seraient fortement critiques vis-à-vis de l'EEE
- Les enseignants à faible estime d'eux-mêmes recevant un feedback positif pourraient être enclins à surestimer les quelques éléments critiques que contient une évaluation globalement positive. Il s'appuie sur les travaux de Moore et Kuol (2005). Bien que conduits sur un faible échantillon (13 enseignants), ils révèlent que certains (5), ont tendance à percevoir les feedbacks des étudiants comme plus critiques qu'ils ne le sont en réalité.
- Les enseignants à faible estime d'eux-mêmes recevant un feedback négatif pourraient être sujets à « l'effondrement » rapporté dans la littérature.

Cette proposition semble importante à explorer dans des études empiriques, à conduire en la référant à d'autres aspects de l'identité professionnelle relatifs à la conception de l'enseignement et de l'évaluation.

2.2. Entre conceptions plurielles de l'enseignement, contextes et trajectoires

La conception de la professionnalité universitaire s'est progressivement élargie d'une focale restreinte sur la discipline à une expertise sur la question de l'enseignement conduisant à l'idée

⁹ Ainsi par exemple le mouvement Slow Science (<http://slow-science.org/>) s'oppose à une vision quantitative et productiviste de la recherche.

d'une culture spécifiée sur l'enseignement et l'apprentissage, *scholarship of teaching and learning* (SoTL), dont les auteurs s'accordent pour trouver l'origine dans les travaux de Boyer (1990) et de la fondation Carnégie (Glassick, Huber, et Maeroff, 1997 ; Kreber et Cranton, 2000), qui ont remis en question la dualité de la recherche et de l'enseignement (avec primat de la première) et proposé de considérer 4 champs d'expertise de l'universitaire (découverte, intégration, application, enseignement), appuyant leur proposition sur des enquêtes à large échelle auprès des universitaires états-uniens. Cette redéfinition de l'universitaire s'élabore en liant les pratiques d'enseignement avec les recherches et les théories éducatives (Weimer et Menges, 1996). Ainsi, l'universitaire se doit non seulement de se tenir au courant des développements dans sa spécialité mais aussi de la théorie, de la recherche et des pratiques sur l'enseignement et l'apprentissage dans sa spécialité. Un tel enseignant, conscient de l'existence et des effets des différentes méthodes, se trouve capable de choisir les plus appropriées et même d'identifier, voire de remettre en questions les présupposés qui gouvernent son travail (Brookfield, 1995). La notion de *scholarship of teaching and learning* désigne également les contributions à un corps de connaissances sur l'enseignement et l'apprentissage (Smith, 2001, p. 53). Elle impose un compte-rendu public d'une partie ou de tout l'acte d'enseignement (concept, conception, actualisation, résultats et analyse), de façon à en rendre possible une revue critique par les pairs et un emploi productif dans le futur travail des membres de la même communauté (Shulman, 1993, p. 6). C'est un changement significatif que de modifier sa conception de l'enseignement, d'une activité qui dérive principalement du travail savant dans sa discipline à une activité informée également par le substantiel et grandissant corps de connaissances sur l'enseignement et l'apprentissage dans le supérieur.

Dans cette évolution, l'évaluation par les étudiants joue un rôle non négligeable si elle participe de l'évaluation formative, concept proposé par Scriven (1967) pour l'amélioration des curricula. L'évaluation formative peut être décrite en termes de source d'information mobilisée (enseignant lui-même, étudiants, collègues, experts), de méthode utilisée pour collecter l'information (questionnaires, observations, entretiens), de degré de temps, d'effort et en termes de modalités adoptées. L'évaluation formative impliquant les étudiants permettrait d'améliorer à la fois le contenu (le corpus de connaissances construit) et le processus (les méthodes utilisées pour le transmettre). On sait que de nombreux enseignants consciencieux améliorent leur enseignement de manière informelle ; ils parlent d'expérience, d'essais/erreurs, d'intuition (Schön, 1983). Parmi les méthodes formalisées, on peut distinguer les « inputs » provenant des étudiants et ceux provenant des collègues ou d'observateurs extérieurs. C'est autour du dialogue et de la réflexion critique que Smith (2001) propose d'organiser les méthodes de recueil d'information sur la pratique. Selon lui, les discussions à l'intérieur d'une communauté de pratiques ont un impact

pour développer et redéfinir la compréhension de l'enseignement. Il se réfère à Brookfield (1995), Glassick, Huber et Maeroff (1997), Kreber et Cranton (2000) ainsi qu'à Cross et Steadman (1996) pour mettre l'accent sur la réflexion critique afin de développer un raisonnement sur la pratique et prendre des décisions mieux informées. L'évaluation de ce qui est, en relation avec ce qui devrait être, qui inclut un feedback conçu pour réduire le fossé, fait partie de ce processus. Le recours à la réflexion autobiographique, aux points de vue des étudiants, aux expériences et perceptions des collègues ainsi qu'à l'étude de la littérature est recommandé. Ces auteurs montrent comment les différentes techniques d'évaluation en cours vont au-delà de la description de ce que les étudiants ont appris et au-delà de l'évaluation de l'impact de l'enseignant. C'est véritablement de la recherche-action qui est préconisée par enquête, analyse des raisons pour lesquelles certaines méthodes fonctionnent mieux avec certains étudiants et élaboration de stratégies alternatives avec contrôle de leur efficacité. Dans la mesure où ces études sont rendues publiques, elles peuvent contribuer à développer une culture experte ciblée sur l'enseignement et l'apprentissage. Celle-ci nécessite une sorte de pensée « méta » par laquelle l'universitaire construit et étudie les questions reliées à l'enseignement-apprentissage (les conditions dans lesquelles il a lieu, à quoi il ressemble, comment l'approfondir...), non seulement pour améliorer son propre enseignement mais pour en faire avancer la pratique générale (Hutchings et Shulman, 1999 ; Hutchings, Huber et Ciccone, 2011). Ajoutons avec McKeachie (2007) que l'EEE peut également être envisagée comme un moyen de faire réfléchir les étudiants à leurs expériences éducatives, de développer des conceptions plus claires des types d'enseignement et des expériences pédagogiques qui contribuent le plus à leur apprentissage (par une meilleure compréhension des buts de la formation), tout en les stimulant à penser de manière plus métacognitive sur leur propre apprentissage et à en accepter la responsabilité. Les investigations conduites par Cross et Steadman (1996) montrent que les enseignants comme les étudiants apprennent à devenir plus efficaces.

Sur le plan de la recherche, ces orientations théoriques et praxéologiques sont à considérer en regard du monde réel de l'enseignement dans ses conditions de milieu, objectives et imaginaires. Les formes d'évaluation sont indissociables des croyances professionnelles, conceptions ou représentations (Pajares 1992) qui sont de plus en plus utilisées pour désigner la manière dont un enseignant conçoit ses pratiques et comment elles évoluent à travers son expérience pédagogique. Les conceptions des enseignants, qui se réfèrent ou s'appuient implicitement ou explicitement sur les diverses modélisations théoriques disponibles ou empiriques traitant de l'apprentissage et de l'enseignement, influencent l'organisation de leur démarche. Ces conceptions tendent à varier selon l'ordre éducatif, la place institutionnelle et les lieux sociaux

dans lesquels les partenaires pédagogiques agissent, en fonction des postures adoptées. Les conceptions à l'œuvre dans les différentes démarches pédagogiques peuvent être considérées comme des connaissances en acte (Vergnaud, 1996) qui tendent à se construire en se référant au parcours biographique ou aux croyances épistémiques des enseignants. L'enseignement est un acte situé, à forte charge émotionnelle. Il s'agit de l'articulation d'un ensemble d'acquis (savoir, savoir-faire et savoir-être) au cours de l'expérience concrète au fil du temps qui guide la pensée et oriente l'action pédagogique nécessitant un travail de bricolage et un ajustement permanent. La manière dont les enseignants conçoivent, ressentent et développent leur démarche pédagogique est aussi liée au contexte institutionnel. Les traditions, les normes et les politiques institutionnelles, les espaces et les dispositifs mis en œuvre par chaque institution au niveau local ou global participent ainsi à influencer leurs conceptions pédagogiques (Becquet et Musselin, 2004 ; Etienne et Annot, 2018 ; Paivandi et Younès, 2019).

L'absence de formation pédagogique initiale et la faiblesse relative du travail collectif au sein des structures de formation peuvent impliquer un bricolage souvent artisanal et personnel dans une construction silencieuse de l'expérience pédagogique enseignante centrée davantage sur des connaissances empiriques et les acquis de ses propres pratiques (Paivandi, 2010). Ainsi les conceptions que développent les enseignants tendent à être intimement liées à leur parcours personnel d'apprenant et d'enseignant. Elles sont également tributaires du contenu ou de l'objet des enseignements dispensés. Assurer un cours magistral et théorique ou un cours pratique, enseigner un savoir, un savoir-faire, un savoir procédural et les usages sociaux de cet objet ne mobilisent pas forcément la même conception de l'enseignement. La spécificité de l'objet visé par l'enseignement est susceptible de devenir l'un des éléments explicatifs du choix pédagogique effectué par chaque enseignant. Il y a aussi la manière dont l'expérience pédagogique se trouve travaillée et réfléchi par l'enseignant : en quoi l'enseignant observe-t-il sa propre pratique et en quoi celle-ci devient-elle l'objet d'une réflexion critique afin de l'améliorer en vue des objectifs visés ?

2.2.1. Typologies des conceptions de l'enseignement

Les travaux de recherche sur les pratiques pédagogiques dans l'enseignement supérieur montrent la pluralité des conceptions (Ramsden, 1992 ; Finkelstein, 1984 ; Entwistle, 1998 ; Fox, 1983). Un travail synthétique qui est réalisé par McKenzie (1995) s'intéresse aux résultats de plusieurs recherches sur les pratiques des enseignants à l'université pour identifier trois catégories de conceptions pédagogiques : les conceptions limitées, les conceptions intermédiaires et les conceptions complètes. Selon l'auteur, les conceptions limitées se réfèrent aux pratiques pédagogiques focalisées sur l'enseignant et son intérêt en tant que détenteur du savoir à

transmettre. Les conceptions intermédiaires se réfèrent aux enseignants préoccupés du contenu et de la façon dont les étudiants s'engagent sur le plan intellectuel et pédagogique permettant d'acquérir le savoir visé par l'enseignement. Enfin, un enseignant tenant une conception complète se concentre sur le développement ou le changement de la conception des étudiants facilitant un apprentissage autonome. Cette posture s'oppose à une transmission unilatérale et à l'imposition du point de vue enseignant. Parmi les autres typologies souvent citées dans la littérature sur la pédagogie universitaire, on peut mentionner les travaux de Trigwell et Prosser (1996) qui se réfèrent à plusieurs recherches pour construire le schéma de cinq conceptions principales. L'approche « A » concerne une stratégie centrée sur l'enseignant cherchant à transmettre les informations aux étudiants; l'approche « B » est centrée sur l'enseignant mais cherche à faire apprendre des concepts et des savoirs aux étudiants; l'approche « C » privilégie l'interaction entre enseignant et étudiant, avec pour objectif l'apprentissage des concepts et savoirs; l'approche « D » s'intéresse à l'étudiant en privilégiant le développement de conceptions par eux-mêmes; l'approche « E » se focalise sur l'étudiant en l'aidant à changer sa conception. Dans une perspective comparative, une recherche internationale, qui a étudié les conceptions de l'enseignement mobilisées par un échantillon d'universitaires de plusieurs pays (Australie, États-Unis, Grande Bretagne, Chine, Singapour, Canada) propose une typologie similaire (Kember and McNaught, 2007). Les cinq catégories identifiées par les auteurs se résument en deux tendances : le modèle centré sur l'enseignant et le savoir, d'une part ; le modèle centré sur l'étudiant et l'apprentissage, d'autre part. Ces catégories varient considérablement selon le type d'enseignement, la discipline, le pays.

Dans le contexte britannique, Entwistle (1998) présente une analyse pessimiste et pense que malgré la pression normative exercée sur les enseignants, ils demeurent souvent attachés aux conceptions traditionnelles de l'enseignement, et que les innovations menées au niveau de leur département ne changent pas d'une manière fondamentale leurs pratiques réelles. De même, il semble que les innovations pédagogiques ne sont pas toujours bien notées car les questionnaires de l'EEE ne permettent pas de les mettre en valeur. En s'appuyant sur les appréciations exprimées par plusieurs groupes d'étudiants, l'auteur montre que la qualité de l'explication, le choix des exemples, les analogies et métaphores, l'enthousiasme de l'enseignant, son empathie, la prise en compte de l'étudiant et son opinion, exercent un impact indéniable sur l'apprentissage étudiant.

Dans une tout autre orientation, Goldin et Adam (2016) se sont intéressés pour leur part aux approches développées par les « excellents professeurs » (lauréats de prix d'enseignement) face aux feedbacks des étudiants. Ils ont exploré cette question via des focus groupes de 2 h. dans une université de recherche en Nouvelle Zélande. Ils ont identifié chez ces enseignants une attitude générale qu'ils qualifient d'approche d'amélioration caractérisée par : une approche réflexive,

cherchant notamment à comprendre le sens des EEE ; une recherche d'amélioration permanente (distincte en cela de l'approche dite du « détecteur de fumée », où l'on approfondit les EEE seulement en cas d'évaluation négative) ; une relation formative à l'EEE considérée non pas comme un jugement de leur personne en tant qu'enseignant mais comme une source d'information utile pour améliorer l'enseignement ; la prise au sérieux des points de vue des étudiants, même quand ils apparaissent inexacts, afin de tenter de trouver des solutions d'amélioration, une approche centrée sur l'apprentissage étudiant, l'EEE étant perçue comme une fenêtre de plus sur cet apprentissage donnant une indication de la mesure dans laquelle les étudiants ont bénéficié de leurs cours ; l'intégration des EEE dans un processus de réflexion complexe amenant à prendre des décisions sur la base d'un processus de jugement ne conduisant pas nécessairement à un ajustement à la demande des étudiants ; un sentiment d'auto efficacité que les enseignants estiment avoir acquis après plusieurs années de pratique.

2.2.2. Réflexivité critique et changements dans le contexte, facteurs de transformation

Selon cette approche réflexive, conceptualisée par Schön (1983), l'enseignant cherche à améliorer ses pratiques par des boucles successives d'expérimentation et d'analyse. Ainsi pour des enseignants réflexifs, quelque soient la qualité de l'EEE ou la valence des feedbacks reçus, ces informations feront de toute façon écho à leur volonté de ménager une place à l'amélioration de leurs pratiques d'enseignement (« room for improvement »). Ces enseignants sont enclins à prendre des risques, à nourrir une pensée divergente, à réfléchir de façon critique et à faire preuve de créativité, et ce dans le but d'améliorer constamment leurs cours. Hendry et al. (2007) rapportent des recherches montrant des liens entre une attitude réflexive et la prise en compte des résultats de l'EEE. Une telle prise de recul permettant d'intégrer le feedback de manière appropriée étant présentée comme un levier du développement professionnel de l'enseignant (Eraut, 2004).

Dans certains cas, le processus pourrait conduire à un changement conceptuel chez l'enseignant entraînant une véritable transformation. Le rôle du feedback consiste ici à bouleverser les convictions de l'enseignant, à lui faire entrevoir des possibilités de régulation, à s'interroger à nouveau sur le sens de ses pratiques, sur la philosophie qui les sous-tend ainsi que sur leur impact sur l'apprentissage des étudiants. C'est ce processus de transformation de l'enseignant que visent Penny et Coe (2004) quand ils affirment que le feedback issu de l'EEE peut, à certaines conditions, amener l'enseignant à penser autrement au sujet de ses enseignements, de ses étudiants et des interactions entre enseignement et apprentissage (p. 244). Ce type d'effet est également à rapprocher des travaux sur le changement conceptuel (conceptual change). Ces travaux, nourris par la psychologie cognitive, ont été initiés par Posner et ses collègues (1982) qui ont étudié le

changement de paradigme au sein d'une communauté scientifique et ont ainsi fondé cette notion de changement conceptuel. Selon eux, celui-ci se produit pour autant que les quatre conditions suivantes soient réunies : [1] une insatisfaction par rapport à la conception existante ; [2] une conception alternative clairement définie ; [3] le caractère praticable de cette conception alternative ; [4] le caractère fécond de cette conception alternative.

Martin et Schwartz (2013) soutiennent que le changement conceptuel s'opère de manière privilégiée quand les exigences du contexte nécessitent le développement de nouvelles connaissances ou l'émergence d'un nouveau cadre intégrateur. Detroz et Verpoorten (2016) associent le changement conceptuel à l'innovation pédagogique : « Pour les tenants de cette théorie [le changement conceptuel], les caractéristiques de l'environnement deviennent donc des facteurs déterminants qui, combinés à des facteurs cognitifs (par ex. la capacité d'un sujet à analyser une situation nouvelle), affectifs (par ex. la capacité à comprendre et gérer ses émotions face à cette situation) et motivationnels (par ex. la valeur perçue), forment le terreau propice (ou non) aux changements conceptuels et à l'innovation. Lorsque le contexte évolue, il peut susciter, avec une amplitude variable - en fonction des caractéristiques individuelles de la personne porteuse de l'innovation -, un déséquilibre cognitif nécessitant un changement conceptuel. Par ailleurs, le rééquilibrage structurant, interne à l'individu et s'exprimant à travers l'action innovante qu'il met en place, trouve dans le contexte les interactions nécessaires à la restructuration conceptuelle ».

2.2.3. Enquête sur les dynamiques d'évolution pédagogique dans deux universités françaises

Nous avons donc cherché à appréhender le lien qui peut exister entre la conception de l'enseignement développée par les enseignants au fil de leur parcours professionnel et la prise en compte des opinions exprimées dans le cadre d'une démarche de l'évaluation de l'enseignement par les étudiants (Paivandi et Younès, 2017 ; 2019). Nous relatons ici de manière détaillée le cadre théorique et les résultats de cette enquête qui mettent en évidence les synergies entre le vouloir et le pouvoir, quant aux dynamiques éventuelles de changement. Ils permettent d'appréhender l'élargissement du paramètre du dispositif et les mises en synergies de divers facteurs qui interagissent, en insistant sur la part du symbolique institutionnel. Il s'agit d'une enquête qualitative par entretiens menés en 2013 et 2014 dans deux universités françaises (université de Lorraine et Université Blaise Pascal à Clermont-Ferrand). Au total, 45 enseignants-chercheurs ont été interviewés dans 4 groupes de disciplines : sciences sociales et humaines (sociologie, sciences de l'éducation, psychologie), sciences expérimentales (mathématiques, biologie, sciences de la terre, chimie), médecine, droit et sciences économiques. Les 45 enseignants-chercheurs

interviewés avaient tous connu l'évaluation de l'enseignement par les enseignants dans au moins un de leurs enseignements assurés durant l'année universitaire en cours ou l'année précédant l'enquête (2012-2013). Le guide d'entretien a été conçu autour de plusieurs thèmes reliant la perception de l'évaluation de l'enseignement par les étudiants et sa prise en compte avec la conception de l'enseignement et de l'apprentissage et le contexte professionnel de l'enseignant. La démarche mobilisée lors de cette enquête a privilégié une approche phénoménologique des processus de changement institutionnel et social car il s'agissait de saisir le point de vue des acteurs concernés à propos du changement institutionnel attendu et effectif ou de la norme mise en œuvre et du sens accordé à leur expérience concrète.

L'hypothèse initiale dans ce travail considérait la conception de l'enseignement comme le facteur explicatif de son attitude à l'égard du feedback étudiant. En effet, derrière chaque conception, il existe un certain nombre de croyances qui tendent à guider la manière dont une démarche pédagogique est conçue et mise en application. L'adhésion de l'enseignant à la philosophie générale de l'EEE est ainsi conditionnée par la conception mobilisée dans sa pédagogie. Les conceptions davantage orientées vers l'apprentissage étudiant tendraient à utiliser d'une manière plus appropriée les résultats de l'EEE en vue de l'amélioration de la pédagogie. Inversement, plus la conception se focalise sur le savoir, moins l'enseignant est enclin à considérer l'étudiant comme un partenaire pédagogique qui a son mot à dire à propos de la démarche pédagogique.

- ***Trois types observés de conception de l'enseignement universitaire***

La première grille d'analyse établie à partir des données des entretiens a permis de constater que l'apprentissage et la manière d'apprendre des étudiants deviennent rarement l'objet d'une réflexion critique chez les enseignants interrogés. L'enseignement est organisé souvent autour d'un savoir à transmettre. La variation des pratiques déclarées se situe entre des méthodes directives et magistrales et des méthodes davantage interactives et participatives. Il s'agit d'un continuum entre deux pôles plus ou moins opposés : le premier privilégie la transmission d'un savoir et une centration sur l'enseignant et le deuxième est focalisé sur une démarche pédagogique améliorée par l'usage des méthodes actives et une centration plus importante sur les étudiants. Ces postures peuvent être résumées en trois types de conception de l'enseignement.

- *Une vision magistrale*

La première conception est focalisée sur le savoir et une vision magistrale de l'enseignement. L'enseignement est essentiellement conçu en tant que transmission d'informations. Les caractéristiques importantes de cette conception sont les suivantes : la finalité de l'enseignement est de transmettre un savoir ; le travail étudiant consiste à assimiler ou à apprendre ce savoir

transmis ; l'acte d'enseigner et l'acte d'apprendre sont séparés ; les étudiants doivent posséder les prérequis et être assez motivés pour apprendre. Pour les enseignants qui adoptent cette vision, la démarche pédagogique n'est pas pensée en lien avec le type d'apprentissage adapté, les besoins ou les pratiques d'étude des étudiants. Certains enseignants disent avoir fait des efforts pour prendre en compte tel ou tel besoin de leurs étudiants, mais l'orientation dominante reste magistrale et transmissive.

La conception focalisée sur l'enseignement et la transmission n'est pas systématiquement une orientation décidée d'une manière unilatérale par l'enseignant. Le type de cours dans le programme et l'organisation pédagogique peuvent lourdement peser sur son choix.

- *La transmission instrumentée*

Selon cette conception, l'enseignement est pensé en tant que contenu organisé et amélioré grâce aux méthodes actives pour faciliter l'apprentissage étudiant. Cette conception est aussi dominée par l'approche transmissive et la finalité recherchée repose sur la maîtrise d'un savoir proposé par l'enseignant mais les enseignants manifestent un vrai souci quant à l'ambiance pédagogique de leurs cours en mobilisant des techniques pédagogiques destinées à faciliter la compréhension et l'accès au savoir, des initiatives visant souvent une meilleure structuration du cours et la participation des étudiants ainsi que l'usage des outils, en particulier, les outils numériques.

- *L'attention portée à l'apprentissage étudiant*

Dans la troisième conception, une attention est davantage portée à l'apprentissage étudiant et sa manière de comprendre et d'apprendre. Dans sa démarche pédagogique, l'enseignant opte pour une structuration signifiante pour les étudiants et s'intéresse à la manière dont ils apprennent et découvrent le signifié du contenu. Cette conception est très souvent mobilisée dans les cours avec un faible nombre d'étudiants, dans les séminaires, les TD ou les TP. La particularité de cette conception repose sur le fait que l'apprentissage étudiant et son intégration au groupe font partie de la réflexion pédagogique de l'enseignant. Le contexte d'apprentissage est pensé en termes de motivation, de ressources et d'accompagnement pédagogique. Il peut s'agir d'activités qui visent directement certains aspects de l'apprentissage étudiant pour lui permettre de mieux appréhender la carte de savoir dans un domaine et de lier les cours entre eux.

Le contexte pédagogique et la structure de l'activité semblent peser sur les conceptions développées par les enseignants. Certains renvoient, au moins partiellement la responsabilité de l'approche mobilisée dans leur cours aux programmes imposés par leur filière. Cette analyse n'est pas totalement fautive car le même enseignant peut mobiliser une conception transmissive dans un cours magistral en licence et une conception améliorée dans un séminaire au niveau master.

Cependant, pour une partie d'entre eux, la vision pédagogique reste plus ou moins invariable selon le type de cours assurés.

- ***Trois types d'attitudes vis-à-vis de l'EEE***

Dans les entretiens, trois attitudes différentes vis-à-vis de l'EEE sont identifiables : la vision négative, la position normative et la position affirmative.

- *Une vision négative*

Une minorité s'oppose ouvertement au principe et à la pertinence de la démarche évaluative effectuée par les étudiants (un entretien sur 5) et mobilise trois types d'arguments pour justifier son opposition : le premier argument avancé cible la pertinence du point de vue de l'étudiant sur l'enseignement. Sont critiqués les aspects normatifs du questionnaire de l'EEE et l'interprétation qui peut être faite des réponses fournies par les étudiants. Cette critique vise aussi les difficultés qu'il y a à formuler une définition unique, acceptée par tous, de ce qu'est un enseignement universitaire de qualité et l'hétérogénéité des attentes individuelles des différents publics étudiants. Il est considéré par ailleurs que la spécificité du métier d'universitaire rend impossible son évaluation notamment par les étudiants. La deuxième raison de l'opposition s'appuie sur l'usage éventuel des résultats à des fins institutionnelles (contrôle, sanctions...). Le troisième souci exprimé par certains enseignants repose sur l'affaiblissement de l'autorité enseignante et une « ingérence » extérieure perçue comme « préjudiciable » au contexte pédagogique.

- *Une vision affirmative*

Les enseignants sont plus nombreux à partager le principe de la pertinence de la prise en compte de l'opinion étudiante afin d'améliorer leurs pratiques pédagogiques. En revanche, ne pas rejeter l'EEE ne signifie pas toujours une forte adhésion à cette démarche. Le degré d'adhésion des enseignants à la démarche de l'EEE varie considérablement selon les cas observés. On repère ainsi l'existence de deux positions parmi les enseignants interviewés qui admettent la finalité formative de cette démarche. La première posture renvoie à une vision affirmative (une forte minorité, près de deux sur cinq), qui adhère d'une manière plus ou moins claire à la philosophie générale de la démarche et à la portée formative de la démarche tout en admettant la pertinence de l'opinion étudiante.

- *Une vision passive-normative*

La posture intermédiaire traduit une vision passive-normative qui accepte l'EEE comme une norme institutionnelle à appliquer sans une adhésion significative à la philosophie de cette démarche (près de deux enseignants sur cinq).

L'appréciation de la démarche de l'EEE est donc aussi influencée par la croyance et la conviction individuelle (valeurs académiques, modèle de l'université, place de la pédagogie...). Les enseignants interviewés qui n'adhèrent pas à cette démarche tendent à justifier leur opposition par la non-pertinence de la démarche et la transformation de l'université dont l'évaluation constitue une composante de cette nouvelle culture contestée. Certains perçoivent aussi l'EEE comme une démarche bureaucratique et sans « âme » mise en place par l'établissement pour répondre aux injonctions institutionnelles. A l'opposé, l'adhésion à la philosophie de l'EEE est parfois justifiée par la critique du fonctionnement actuel de l'université et le souhait d'une université davantage préoccupée par les questions pédagogiques et le devenir des étudiants en difficulté.

- ***Effets de l'EEE sur les pratiques pédagogiques***

Les attitudes générales des enseignants à l'égard de l'EEE tendent à se référer à la fois à une problématique personnelle s'exprimant dans la manière dont chaque enseignant conçoit la mission d'enseignement, et à une logique institutionnelle qui impose sa norme à tout le monde. Le mode d'articulation dans la perception de l'enseignant de cette double entrée est une source importante de compréhension du discours enseignant.

Les enseignants interviewés qui ne se déclarent pas « hostiles » à la démarche de l'EEE ou qui sont plus ou moins favorables à cette démarche s'intéressent plus souvent aux appréciations des étudiants et sont sensibles à ce feedback. On peut dire que ces enseignants s'approprient la dimension formative de cette démarche et, d'une manière réfléchie et active, tendent à la mettre au service de leur développement professionnel sur le plan pédagogique. Les informations recueillies sont ainsi susceptibles de devenir un facteur de changement et d'amélioration des enseignements et un outil formatif pour le développement professionnel. Cet enjeu trouve tout son sens lorsque l'EEE est bien appréciée par les enseignants notamment avec un soutien institutionnel éventuel. Le contraire interroge l'enseignant sur l'utilité de la démarche. Pour donner une réponse précise à la question liée à l'impact réel de l'EEE sur les pratiques pédagogiques, il a fallu voir en quoi les enseignants tentent de mieux appréhender les questions liées à l'acte d'apprendre des étudiants et à leurs difficultés d'apprentissage, et comment ils peuvent faire évoluer éventuellement leur conception de l'enseignement.

Les données des entretiens montrent qu'accepter la logique de l'EEE ou adhérer à sa philosophie ne veut pas dire pouvoir ou vouloir changer les différents aspects de ses pratiques pédagogiques. Selon les enseignants, les critiques ou les demandes exprimées par les étudiants sont très différentes, allant du contenu des cours à la présentation du cours ou aux améliorations liées à

l'accès aux documents. En effet, face aux multiples demandes formulées par les étudiants, selon les cas expliqués par les interviewés, quatre types de réaction ont été identifiés :

- *Refus* : le refus de l'ajustement peut reposer sur la non-pertinence de la demande étudiante selon l'enseignant et renvoie souvent aux aspects « non-négociables » du cours pour l'enseignant. Le feedback « négatif » concerne un aspect dont on attribue la responsabilité aux étudiants et on ne pense pas pouvoir agir sur celui-ci. Certaines demandes sont considérées revêtir un caractère « insensé », « peu sérieux », « impossible » selon les enseignants. Ce type d'interprétation s'appuie sur l'expérience concrète de l'enseignant et ne représente pas une attitude systématique.

- *Impossibilité de changer* : le refus de mettre en cause la démarche pédagogique et d'envisager les ajustements demandés ne sont pas systématiquement liés à une opposition à l'EEE. Le feedback « négatif » peut être attribué au curriculum qu'il n'est pas possible de changer ou d'influencer. Lorsque le contenu du cours est critiqué, les enseignants n'ont pas toujours la possibilité de le réajuster ou de le modifier même s'ils sont conscients de l'importance de la demande étudiante. Certains ajustements demandés par les étudiants ne sont pas réalisables selon le plan matériel. Les enseignants sont conscients de la pertinence des propositions avancées par les étudiants sans pour autant pouvoir modifier la structure pédagogique de leur enseignement.

- *Changement marginal* : Le changement marginal concerne les enseignants qui disent avoir réajusté des aspects plutôt « techniques » de leur pédagogie sans grande importance. En fait, lorsque des remarques faites dans une EEE visent des ajustements techniques marginaux comme « *mettre les documents sur la plateforme* », « *changer le rythme* » ... les enseignants tendent à les intégrer plus facilement. La grande majorité d'enseignants rencontrés (même quand ils sont opposés à l'EEE) accepte ce type de modification.

- *Révision de la démarche pédagogique* : Certains enseignants acceptent de revoir leur démarche pédagogique en prenant appui sur le feedback étudiant. Cette réflexion pédagogique ouverte concerne essentiellement les éléments constitutifs de la démarche pédagogique et semble exclure certains champs perçus comme « non-négociables ». Mais la réflexion critique sur l'enseignement et la démarche pédagogique consistant à s'appuyer sur les résultats de l'EEE afin de développer une critique substantielle de la démarche pédagogique constitue un phénomène rare.

- ***Variations entre vouloir changer et pouvoir changer***

Les enseignants se trouvent parfois face aux contradictions inhérentes à l'EEE quant aux opinions de leurs étudiants. Lorsqu'une très grande majorité d'étudiants critique tel ou tel aspect de la

pédagogie ou font des remarques plus ou moins convergentes, la réflexion pédagogique semble poser moins de problème. Réfléchir sur le changement tend à devenir plus compliqué lorsque les étudiants sont assez partagés sur tel ou tel aspect de la démarche pédagogique et expriment des opinions contradictoires. Cette situation est tout à fait compréhensible lorsque les étudiants n'ont pas les mêmes besoins en matière d'apprentissage ou d'encadrement pédagogique. Lorsque les enseignants ne s'opposent pas d'une manière radicale à l'EEE, on peut se trouver face à plusieurs configurations possibles à propos de la posture enseignante :

- Le premier cas concerne les enseignants pour qui coïncident l'attitude « vouloir changer » et la possibilité effective de « pouvoir changer ».
- Le deuxième cas concerne les enseignants pour qui malgré l'existence d'un « vouloir changer », la prise en compte n'est pas envisageable à cause de la nature du changement demandé (contraintes institutionnelles, matérielles ou pratiques).
- Le troisième cas comprend les enseignants qui refusent la pertinence de la demande étudiante ou de l'EEE (ne pas vouloir changer) malgré la possibilité de changer la démarche pédagogique (pouvoir changer) ;
- Le quatrième cas renvoie aux enseignants défavorables à l'EEE, qui sont face à une demande irréalisable. Le caractère irréalisable de la demande étudiante est susceptible de devenir un alibi pour rejeter l'EEE comme démarche formative au service de la pédagogie à l'université.

L'impact de l'EEE sur les pratiques pédagogiques est souvent tributaire de la position de l'enseignant vis-à-vis de cette démarche. Une opinion négative à l'égard de l'EEE diminue la possibilité de faire évoluer la pédagogie à partir du feedback étudiant sans pour autant le bloquer entièrement. Une vision négative de l'EEE ne bloque pas définitivement les effets éventuels sur la démarche pédagogique. Cependant, son impact demeure limité et va rarement très loin.

- ***Variations entre contextes, conceptions et trajectoires***

La conduite des enseignants vis-à-vis de leur pédagogie est aussi influencée par un ensemble de facteurs liés à leur parcours, à la façon de concevoir leur carrière universitaire, à leurs stratégies personnelles ainsi qu'au contexte disciplinaire. Chaque contexte disciplinaire constitue un sous-univers possédant un *style d'existence* (Goffman, 1991). Pour Musselin (2008) comme pour Clark (1987), la discipline et le type d'établissement sont déterminants dans les modalités de l'activité d'enseignement comme de recherche. Dans l'ensemble de l'enquête, la reconnaissance de la pertinence de la démarche de l'EEE est plus marquée dans les disciplines scientifiques. Par exemple, dans une filière scientifique, le débat autour de la baisse des effectifs a conduit à une mobilisation collective et à la prise en compte des résultats de l'EEE pour faire évoluer la

démarche pédagogique. Un autre exemple concerne un département en SHS qui suite à une mauvaise expérience de l'EEE (commentaires très négatifs et nominatifs visant certains enseignants) est devenu réticent à l'égard de cette démarche.

Il est à noter que les maîtres de conférences et les professeurs n'investissent pas la pédagogie de la même manière au début de leur parcours, durant les moments critiques de leur parcours ou vers la fin de leur parcours. Un maître de conférences qui n'a pas l'intention de passer une HDR pour devenir professeur peut développer une stratégie alternative en s'intéressant davantage aux questions pédagogiques.

La conception de l'enseignement tend souvent à exercer un impact direct sur la manière dont l'opinion étudiante est prise en compte par les enseignants. Les enseignants qui se trouvent dans une perspective transmissive tendent à être moins ouverts aux ajustements de leur démarche en prenant en compte les résultats de l'EEE. Certains acceptent d'apporter des modifications qui portent essentiellement sur les aspects organisationnels ou matériels de leur cours. A l'opposé, plus l'enseignant focalise sa démarche pédagogique sur l'apprentissage étudiant, plus il a tendance à s'intéresser à l'opinion étudiante afin d'améliorer sa pédagogie. En effet, les enseignants de cette catégorie mobilisent une palette plus variée de méthodes pédagogiques pour animer leurs cours et sont plus ouverts à une réflexion critique sur la pédagogie. Le feedback étudiant devient ainsi une ressource importante pour s'intéresser au développement professionnel et améliorer la pédagogie.

Le lien observé entre la conception de l'enseignement et le fait de tenir compte des résultats de l'EEE pour réajuster ou faire évoluer sa pédagogie montre que l'impact effectif de l'EEE sur les pratiques pédagogiques n'est pas un phénomène mécanique et linéaire. Les enseignants qui sont dans une logique transmissive ou qui assurent un enseignement magistral (dans le programme) sont nettement moins disposés à changer leur démarche pédagogique.

De manière générale, la relation entre la dimension formative de la démarche de l'EEE, son appréciation par les enseignants et la conception de l'enseignement n'est pas linéaire et figée. Il s'agit d'une dynamique productive (et non pas reproductive) car un impact positif ou négatif de l'EEE est susceptible de créer une réflexion critique, même limitée, à l'égard de la conception ou vis-à-vis de l'appréciation de l'EEE. Le processus d'une réflexion critique sur la pédagogie à l'université ne va pas de soi, même lorsque les enseignants ont la volonté de faire évoluer leur pédagogie afin d'accompagner mieux les étudiants. L'enquête a montré que les pratiques pédagogiques décrites par les enseignants interviewés s'inscrivent majoritairement dans un paradigme de l'enseignement transmissif traditionnel. D'où une vraie interrogation par rapport à la portée formative possible de l'EEE dont les effets produits souvent réels restent cependant

limités. Sans une formation, un accompagnement et une réflexion pédagogique critique continue individuelle et collective, une démarche comme l'EEE semble être prisonnière de ses limites. Un tel modèle pédagogique s'avère prégnant malgré des variations liées aux facteurs culturels, à la fois institutionnels et disciplinaires.

2.2.4. Perspective élargie de l'EEE

L'EEE n'intervient pas dans un vide social. Elle y introduit une nouvelle norme et une pratique visant à transformer l'ordre pédagogique habituel. Les normes affectent les individus qui discutent de leur pertinence à partir ce que Castoriadis (1975) avait appelé les « significations imaginaires » d'une institution dotée d'une histoire forte et d'un ensemble de traditions. Il s'agit donc d'appréhender l'esprit dans lequel une norme comme l'EEE émerge et la façon dont celle-ci se conjugue avec les mœurs (Paivandi et Younès, 2017). La démarche peut contribuer à la transformation d'un système social ou plus exactement dans le contexte professionnel d'un système socio-mental, d'un changement intentionnel voulu par l'institution. En introduisant cette norme, l'institution tend à développer « la perspective du changement d'attitude de valeurs » (Lewin, 1959) des enseignants. En tant que démarche institutionnelle, l'EEE qui n'a pas été créée par l'enseignant, ou par son groupe d'appartenance (département), provient d'une source extérieure à l'univers pédagogique. Cependant, dans sa visée formative, elle cible l'ajustement pédagogique et le travail enseignant. Considérer l'EEE comme une norme mise en place dans le cadre d'une institution implique que l'on ne l'analyse, comme le dit Garfinkel, pas pour elle-même, mais quant aux propriétés des événements qu'elle engendre (Coulon, 1993, p. 199), ce qui est un élément théorique central pour l'analyse de l'évaluation et ses effets invisibles éventuels. Au sein d'une institution, les normes et les règles sont l'objet de négociations permanentes entre les acteurs, elles ne peuvent pas être définies par l'institution d'une manière unilatérale. Les enseignants utilisent les « règles préférentielles » qui respectent sans doute la procédure de base, mais qui sont perçues comme meilleures et appropriées par eux selon la conception qu'ils ont de la mission d'enseignement. Autrement dit, dans le cadre d'une analyse sociologique, il faut pouvoir distinguer entre les normes institutionnelles (la mise en place de la démarche de l'EEE dans le cadre d'une stratégie locale au niveau de chaque établissement, par exemple) et les actions des enseignants concernés. La façon dont l'EEE est perçue et appréciée, comme n'importe quel phénomène social, n'est pas un reflet fidèle du cadre normatif et institutionnel. Les enseignants préservent une autonomie relative d'action et les normes institutionnelles comme l'EEE n'existent qu'à travers une action filtrante et « déformatrice » des situations locales et des acteurs. L'enseignant en tant qu'acteur n'aborde pas son institution comme agissant exclusivement selon un système de normes. Son action est également définie par ses valeurs,

l'environnement pédagogique de sa filière ainsi que les relations qu'il noue avec autrui (étudiants et collègues), contribuant à sa mobilisation.

L'enquête s'est aussi intéressée à la variation des attitudes adoptées par les enseignants selon leur conception de l'enseignement et en évitant d'examiner le corps enseignant déconnecté de son environnement professionnel et de la société, mais en tentant d'appréhender l'économie des affects et appréciations individuels à partir de l'intériorisation des normes universitaires. En effet, la démarche de l'EEE n'est pas neutre et sa mise en place subit une altération dans chaque contexte universitaire. Le contexte, l'origine et l'enjeu de l'EEE tendent à participer à définir les conditions de son application. L'élément crucial semble se manifester au niveau de la perception d'une norme instituante comme l'EEE, susceptible de faire surgir des tensions ouvertes ou latentes. Une tension apparaît alors si sont mis en cause les codes et les pratiques « habituels » ou une tradition perçue comme légitime et pertinente.

Des critères et des facteurs explicatifs des différences ou des similitudes observées dans la prise en compte du feedback des étudiants dans les pratiques pédagogiques des enseignants ont été dégagés afin d'examiner les effets réels produits par l'EEE dans le contexte de plusieurs disciplines universitaires. Il a été tenté d'appréhender la composante objective et fonctionnelle de l'évaluation de l'enseignement et la composante subjective et imaginaire de la démarche. En travaillant sur le discours enseignant, on est appelé à se référer à l'usage des significations, tout en les examinant dans leur rapport avec l'activité sociale et le contexte local disciplinaire.

Le choix de l'approche qualitative pour saisir le point de vue des acteurs et le sens attribué à leur expérience a conduit à écarter l'usage des opérations chiffrées. Cette inscription méthodologique a impliqué de privilégier la signification des attitudes et des opinions sans être dans une logique de représentativité numérique, de vérification, de statistiques et de procédures codifiées.

Les préoccupations de la recherche sur l'EEE d'abord essentiellement d'ordre psychométrique ont progressivement intégré des questionnements pédagogiques portant sur les effets et les variables médiatrices de ces effets qui s'envisagent aux différents niveaux de la culture, des dispositifs et des expériences, en soulignant les interdépendances. Car si les dispositifs sont des incarnations de la culture, ils participent également à sa transformation. De même la culture façonne l'expérience de l'EEE dans les significations qui lui sont attribuées en même temps que l'expérience de l'EEE façonne la culture de l'enseignement. L'expérience de l'EEE est certes fonction des dispositifs mais c'est la manière d'en faire l'expérience qui leur donne du sens. Ces réciprocités productrices de sens et de tension entraînent de redoutables questions épistémologiques et méthodologiques quant aux dispositifs mis en œuvre et comment étudier

des effets quand aucun critère simple ne peut réduire la complexité et aucun modèle causaliste ne peut permettre de l'appréhender.

Partie 3 – Les dispositifs d'évaluation comme médiations psycho-socio-techniques

La mise en perspective des travaux conduits dans les champs de l'évaluation de l'enseignement et des apprentissages des étudiants participe d'une approche critique des dispositifs d'évaluation conçus comme médiations psycho-socio-techniques. Les dispositifs d'évaluation sont appréhendés dans une orientation interprétative et critique dans la lignée de Stake (1975, 1976), Guba et Lincoln (1989), Lincoln et Guba (1985), Patton (1980), Cranton (2001), soulignant l'enjeu d'une articulation des conditions de milieu et de sens, plutôt que d'en rester à une quantification et objectivation (Younès et Romainville, 2012 ; Younès et Gay, 2014). A travers ce paradigme interprétatif et critique, on peut réinterroger une tradition de recherche sur l'évaluation portant sur la fiabilité et la validité des mesures quantitatives versus qualitatives. Cette approche ne représente qu'une perspective collective, où les voix individuelles sont additionnées et « fondues » dans une moyenne. Dans une approche quantitative, ce sont les différences individuelles qui sont minorées au profit des tendances majoritaires.

En ce qui concerne les EEE, Cranton (2001, pp. 15-16) souligne que bien qu'elles conduisent à des résultats statistiques, elles sont toujours subjectives et interprétatives. La crainte que les résultats qualitatifs ne soient qu'un ensemble d'opinions aléatoires, non généralisables et non représentatives est la critique communément faite à la recherche interprétative. Au contraire, Cranton ne considère pas que la subjectivité soit à dépasser et refuse de parler de « biais », étant donné que selon les contextes, les caractéristiques d'un enseignement approprié sont variables et que l'évaluation des étudiants se fait à travers le filtre de leurs valeurs et conceptions d'un bon enseignement.

Dans un paradigme qualitatif, la profondeur et la signifiante sont privilégiées. Certes, les résultats de ce genre d'évaluation ne sont pas concis comme le sont des synthèses statistiques de résultats quantitatifs, même si rien n'empêche de les présenter dans des formats faciles à lire. Toutefois, des résumés, des citations et un style narratif peuvent se révéler plus pertinents et informatifs que des tableaux et des graphiques. Par ailleurs, le rapport personnel ou l'auto-évaluation sont des composants centraux de l'évaluation interprétative. L'enseignant peut y décrire la philosophie de sa pratique, faire des commentaires sur les perceptions des étudiants ou souligner son plan de développement. C'est d'ailleurs la philosophie des « dossiers d'enseignement » via lesquels les enseignants sont invités par certains établissements à rassembler, contextualiser et interpréter les résultats de leurs EEE.

Deux dispositifs d'évaluation significatifs comme médiations psycho-socio-techniques sont présentés. Tout d'abord, la conception d'un dispositif d'EEE informatisé et l'étude de ses effets. Ensuite, un outil d'évaluation interactive (boitiers de vote) utilisé dans le cadre d'un dispositif de recherche/accompagnement.

3.1. La conception et l'opérationnalisation d'un dispositif d'EEE informatisé et l'étude de ses effets

Comme nous venons de le rappeler, l'évaluation de l'enseignement se trouve souvent limitée par un cadre d'interrogation organisé habituellement autour d'un nombre fini de critères prédéterminés et par un traitement de l'information principalement quantitatif. Le recours aux technologies tant pour le recueil des données que pour leur traitement exacerbe cette approche normative et généralisante dont les effets déformants et les dangers ne sont pas à minimiser. Dans le contexte de la mondialisation, un large accord existe désormais quant à la nécessité de l'évaluation de l'enseignement ainsi que de modélisations et de référentiels génériques dont le grand avantage est qu'ils permettent de s'entendre sur des fondamentaux. Par exemple, le processus de Bologne, qui marque la volonté de construire l'espace européen de l'enseignement supérieur, entraîne une convergence des systèmes universitaires avec l'adoption d'une architecture des diplômes comparable et avec l'adoption également du système européen de crédits dits « ECTS » visant la compatibilité des systèmes d'enseignement supérieur. Il s'agit aussi d'assurer la reconnaissance des diplômes au niveau européen et de développer des méthodologies d'évaluation de la qualité pour instaurer des mécanismes d'évaluation et d'habilitation mutuellement acceptés. Le référentiel européen d'assurance qualité (ENQA, 2005) a donc été retenu par l'ensemble des pays et des agences nationales d'évaluation (voir par exemple pour la Belgique, Duykaerts et Malmedy, 2014).

Les analyses montrent cependant que ce référentiel commun donne lieu à des démarches d'évaluation très différentes (Charlier, Croché, Fallon et Leclerc, 2014), conduisant ou non à des synthèses quantitatives sous forme de notation et de classement suivant les contextes nationaux. Si l'on prend l'EEE, elle s'est largement développée autour de la construction de référentiels génériques de la qualité de l'enseignement. Elle repose généralement sur des questionnaires fondés sur ces référentiels, avec l'hypothèse implicite que les items correspondent à des critères universels et transversaux aux filières et programmes de formation. Certes, il existe un large accord en ce qui concerne le caractère multidimensionnel de l'enseignement et des généralisations ont pu être établies à partir de recherches empiriques aboutissant à des dimensions validées (Marsh, 1982, 1987) quant à leur impact sur la performance pédagogique :

enthousiasme du professeur, structuration et étendue du cours, attentes élevées vis-à-vis des étudiants (difficulté du cours, travail demandé), clarté des attentes et des modalités d'évaluation des étudiants, qualité des interactions suscitées entre les étudiants, relation positive entre l'enseignant et l'étudiant, qualité des apprentissages. Cependant les chercheurs ne sont pas tous d'accord sur le nombre de dimensions structurantes (28 selon Feldman, 2007 ; 4 selon Abrami, d'Apollonia et Rosenfield, 2007) et d'aucuns considèrent que certaines d'entre elles sont plus influentes. De plus, les dimensions retenues n'ont pas toutes un effet prouvé sur la qualité des apprentissages. Par exemple, l'estimation du caractère communicatif de l'enseignant ou celle de la difficulté du cours ne prédisent pas la réussite des étudiants (Abrami, d'Apollonia et Rosenfield, 2007).

Par ailleurs, ces dimensions de performance ont également été discutées quant à leur adaptation aux différents types d'enseignement. Ainsi il se peut que, dans des modalités particulières, les dimensions mentionnées ci-dessus telles que l'organisation du cours ou l'enthousiasme de l'enseignant se révèlent inappropriées. Si la plupart des étudiants sont sécurisés et ont besoin d'un haut degré de structuration (Rodger, Murray et Cummings, 2007), celle-ci peut en étouffer d'autres, être contreproductive par rapport aux objectifs ou encore être inadaptée à certaines méthodes pédagogiques (Mc Keachie, 2007). Kolitch et Dean (1999) ont réalisé une analyse critique du questionnaire utilisé à l'Université d'état de New York du point de vue du modèle « transmissif » ou « centré sur l'enseignement » relativement au modèle de l'enseignement « critique-engagé ». Ils en ont conclu que la majorité des items se rapportait à un modèle de l'enseignement transmissif, développant ainsi une vision étriquée de ce qu'est un cours « efficace ». Ils ont fourni des exemples d'items alternatifs reflétant un modèle de l'enseignement « critique-engagé ». Abrami, Theall et Mets (2001) ont établi des constats similaires, par exemple à propos de l'apprentissage coopératif et ont proposé des items ayant trait à la capacité de l'enseignement de faciliter une interdépendance positive (par exemple : les étudiants étaient-ils responsables de l'apprentissage de leurs pairs ? ». Ils ont fait de même à propos de l'apprentissage auto-régulé, de l'approche par problème... Dans de nombreux cas, il apparaît donc que les questionnaires sont basés sur un modèle de l'enseignement traditionnel, transmissif, inadapté à certains types de cours et pouvant même participer à contrecarrer l'évolution pédagogique.

En ce qui concerne la qualité des apprentissages dont nul ne peut remettre en cause qu'il s'agisse d'un objectif majeur (si ce n'est l'objectif principal) d'un enseignement, deux critiques sont apportées selon un ordre de raison tout à fait différent. Abrami remarque à juste titre que la plupart des formulaires d'évaluation utilisés sont très évasifs sur ce point qu'ils abordent de

manière très générale (par exemple : « comment jugez-vous les apprentissages liés à ce cours ? ») et propose de les spécifier (Abrami, d'Apollonia et Rosenfield, 2007). Notons cependant que certains questionnaires offrent une investigation plus poussée à ce niveau. Par exemple le formulaire de l'IDEA révisé (Hoyt et Lee, 2002) distingue 12 types d'apprentissages (connaissances factuelles, résolution de problèmes, expression orale, travail en équipe...). La seconde critique renvoie au caractère imprévisible des effets de l'enseignement (Leclerc, 2012), à savoir la difficulté d'identifier ce qui va agir, à quel moment et chez qui.

Il semble donc que l'existence de principes généraux d'un enseignement efficace puisse être remise en question et qu'il ne s'agit, à tout le moins, pas de vérités absolues et invariables, quand bien même seraient dégagés des éléments de base communément admis et partagés. L'intérêt des référentiels de critères transversaux est néanmoins double. Ils permettent d'identifier des critères communs et d'établir des procédures techniques standardisées facilitant des comparaisons. L'ampleur de la révolution numérique va d'ailleurs accentuer ces tendances, au point d'exercer une fascination qui risque d'être aveuglante.

3.1.1. A propos des dispositifs d'informatisation de l'évaluation

L'informatisation des systèmes d'évaluation et leur utilisation sur Internet leur confèrent des avantages incontestables du point de vue de l'opérationnalisation de la démarche d'EEE (Theall et Franklin, 2001 ; Younès, 2005) :

- l'informatisation permet de mettre immédiatement à disposition les questionnaires n'importe où dans le monde et de les gérer aisément dans leurs évolutions ; ainsi les modifications effectuées au niveau central sur le serveur sont instantanément répercutées sur le terrain ;
- à partir de chaque ordinateur connecté, il est possible suivant un profil défini de saisir les réponses aux questionnaires ou de consulter les résultats accessibles à tout moment ;
- les résultats étant stockés dans une base de données, leurs traitements sont automatisés et conservés, résolvant ainsi les problèmes de collecte, de dépouillement et de stockage ;
- des indicateurs et des tableaux de bord peuvent être édités, ce qui favorise la visualisation des résultats de l'EEE et même la comparaison des enseignements, des programmes, des composantes.

Toute une série d'articles ont été publiés quant aux avantages et aux inconvénients de la passation des questionnaires en ligne ou papier (Anderson, Cain et Bird, 2005 ; Avery, Bryant, Mathios, Kang et Bell, 2006 ; Norris et Conn, 2005 ; Nulty, 2008). Il ressort des enquêtes menées à ce sujet que les taux de réponse aux questionnaires en ligne sont inférieurs à ceux sous forme de papier, ce qui conduit certaines universités à préférer conserver la modalité traditionnelle. Il nous

semble toutefois que les conclusions tirées de ces enquêtes confondent l'outil et les conditions de passation. C'est parce que les étudiants sont captifs qu'est observé un taux supérieur de réponses dans le cas où les questionnaires sont distribués et récupérés séance tenante. Des taux de réponse tout à fait importants lors de passation en ligne ont pu être observés quand des conditions d'engagement des étudiants étaient créées, que ce soit par imposition ou par un processus d'implication (Younès, 2006).

Figari et Remaud (2014) soulignent l'importance de faire du « sur-mesure » négocié, même dans le cas de référentiels « déjà- là ». Postiaux (2010) montre notamment comment, dans l'enseignement supérieur, les référentiels de formation spécifiques construits en négociation par des équipes d'enseignants, et ce dans différents pays, sont des cadres utiles pour le pilotage de la formation, les pratiques d'enseignement et d'évaluation ainsi que pour la construction d'une identité collective située. Une voie importante se dégage donc amenant à repenser l'EEE entre généralité et spécificités contextuelles.

Dans notre thèse intitulée *L'effet évaluation de l'enseignement supérieur par les étudiants*, nous avons choisi de questionner la difficile implantation de l'évaluation de l'enseignement dans le système universitaire français en ciblant plus particulièrement les effets de l'introduction dans une université française d'un dispositif informatisé d'évaluation de l'enseignement par les étudiants selon la méthodologie de la référentialisation (Figari, 1994, 2001, 2014). Cette forme d'évaluation est particulièrement significative puisqu'elle touche au cœur de l'acte pédagogique et qu'elle se situe à la confluence de trois principaux acteurs du système d'enseignement universitaire (les étudiants, les enseignants et les responsables). Elle nous a conduit à questionner à la fois l'enseignement et le système d'enseignement. Il était plus particulièrement cherché à appréhender les effets de l'implantation d'un dispositif informatisé d'EEE à visée formative, considérant en quoi et comment l'EEE peut constituer un seuil critique de passage d'un système d'enseignement à un autre. Nous n'avons pas appréhendé ce seuil en interrogeant uniquement une catégorie d'acteurs, mais en considérant, dans une vision systémique les interactions entre acteurs confrontés à des dispositifs qui les réunissaient (ou les séparaient). L'étude a porté sur le positionnement de l'institution universitaire et des acteurs vis à vis de l'EEE, à partir de la participation à l'évaluation, des informations issues des EEE et des effets perçus et escomptés par les acteurs. La notion d'effets de seuil a été pensée non comme une limite clairement déterminée mais comme une suite de passages pouvant conduire à modifier profondément les caractéristiques du système d'enseignement. Les résultats de l'investigation de l'impact de l'introduction du dispositif d'EEE ont mis en évidence, de manière transversale, l'ambiguïté et les paradoxes de ses effets dans la remise en cause du système d'enseignement traditionnel. Mais il a

également été analysé que l'implantation d'un tel dispositif à visée formative n'a pas suffi à établir le passage d'un système d'enseignement S1 (basé sur un enseignement individuel dans lequel les méthodes d'enseignement ainsi que les modalités d'évaluation des étudiants ne sont pas nécessairement explicitées) à un autre système d'enseignement S2 supposant au contraire d'explicitier les visées et méthodes pédagogiques, de rendre compte de ce qui se passe dans l'enseignement et de favoriser une démarche collective. Ce qui nous a conduit à envisager les limites de ce dispositif d'évaluation. Il est apparu que l'informatisation du dispositif et sa mise sur Internet avait accentué les effets de l'EEE et que des modalités d'EEE adaptées à la culture universitaire et la construction de dispositifs stimulants conservant un aspect évolutif restaient à inventer. Le référentiel doit pouvoir être transformé sous peine d'enfermer les pratiques dans des cadres fixés contrevenant mêmes aux objectifs et aux dynamiques pédagogiques. C'est la prise en compte des écarts par rapport à une norme provisoirement stabilisée qui donne au référentiel son caractère évolutif ajusté aux pratiques effectives.

La conception et la mise en place du dispositif d'EEE s'ancrait dans une approche de l'évaluation comme formative, à savoir comme processus de régulation subjectif et intersubjectif inhérent au développement et à l'évolution de l'action pédagogique (Stufflebeam et al., 1980). La référentialisation en tant que processus assorti de procédures critériées (Figari, 1994) en constituait l'opérateur. Cependant, elle n'a pu qu'être démarrée (voire empêchée) étant données les conditions de milieu et notamment les résistances du terrain au niveau de certains enseignants avec lesquels il n'a pas été possible d'enclencher un processus de coconstruction d'un sens en partage. Par contre a pu être mis en place un logiciel d'évaluation de l'enseignement fonctionnant sur Internet qui a constitué une première médiation psycho-socio-technique opératoire à l'Université d'Auvergne. Ce logiciel s'est révélé un élément clé aussi bien pour générer une dynamique institutionnelle située que comme moyen de régulation. Son accessibilité a rendu possible une participation de tous les étudiants et une consultation aisée des résultats. Son adaptabilité correspondait à une évaluation contextualisée et évolutive. Ce dispositif combinant des jugements critiques individuels qualitatifs et des indicateurs basés sur des statistiques, a permis d'établir des comparaisons sur un certain nombre de critères sélectionnés tout en conservant la richesse liée aux variations inter individuelles. Il a fourni de nombreuses informations sur l'enseignement, la formation, les étudiants et les enseignants. Il a également contribué à passer de l'implicite à une forme d'explicitation quant à l'enseignement, aussi bien en termes de visées que de modalités. On peut considérer que la première mise en œuvre, bien que portée par un nombre réduit d'acteurs, a néanmoins eu un effet d'enclenchement à partir duquel d'autres acteurs ont été progressivement engagés. La réflexion sur les objectifs, les critères et

indicateurs, les modalités de la formation comme des enseignements et leur adaptation au public a pu se développer graduellement à partir de l'explicitation.

Un tel dispositif informatisé d'EEE qui a ainsi établi une médiation contextualisée et opérante du développement du processus d'évaluation confirme l'importance d'intégrer les trois facteurs suivants :

- une volonté institutionnelle des décideurs de l'université d'amorcer l'EEE, qui ont relayé la politique nationale en définissant un cadre approprié au contexte local et des moyens pour l'évaluation ;
- une participation des acteurs concernés à tous les niveaux de la conception et de la mise en œuvre ;
- une interface informatisée servant de médiation entre les acteurs impliqués dans le système d'enseignement (étudiants, enseignants et responsables).

L'efficacité de ce dispositif pour la collecte, le traitement et la mise à disposition de l'information en a fait un outil adapté à une évaluation-régulation qui nécessite une instrumentation permettant un feedback souple et rapide sur les activités d'enseignement. De telles pratiques présentent un intérêt particulier en tant que premier moment d'une dynamique de régulation (Hadji, 1992) qui a pour fin de procéder à des ajustements, par l'utilisation adéquate de l'information en retour (Allal, 1988), le point de vue des étudiants contribuant à une dynamique de formation plus partenariale. Un autre atout d'une telle médiation est sa capacité à évoluer dans le temps en fonction des transformations d'un milieu de formation en permettant des informations renouvelées et souples sur les évolutions du milieu et du dispositif d'évaluation.

Cependant, comme toute technique, une telle évolution n'est pas sans créer de véritables risques : une dérive progressive vers une bureaucratisation du processus, à savoir vers un dispositif plus formel que réellement créateur d'une dynamique collective autour de la qualité de la formation, voire une aliénation des étudiants comme des enseignants à une logique consumériste de « client » appréciant les services rendus par un prestataire. Au regard de certaines pratiques - en cours dans les pays dont les systèmes d'enseignement sont les plus libéraux - qui consistent à utiliser les résultats des EEE pour recruter, titulariser ou promouvoir les enseignants, il n'est pas déraisonnable de considérer qu'à l'extrême de cette logique, ceux-ci pourraient en être réduits à une position de prestataires de services plus ou moins sommés de satisfaire les étudiants sous peine de sanction. Tout un courant développe par ailleurs une approche extrêmement critique quant à l'EEE en tant que telle (Sproule et Vâlsan, 2009). A la menace identitaire pour les enseignants s'ajoute une interrogation sur la place accordée aux étudiants. Demander à l'étudiant d'exprimer un avis critique sur l'enseignement qu'il a reçu, c'est lui conférer un pouvoir

d'expression et une place active. Cependant, la procédure d'EEE la plus courante, reposant sur un questionnaire anonyme en fin de cours, s'avère placer l'étudiant dans un processus prédéterminé, dans lequel le sens de l'expérience pédagogique et la confrontation des points de vue peuvent être court-circuités.

En outre, de manière paradoxale, la part de l'engagement de l'étudiant dans la qualité de l'enseignement se trouve évacuée du fait qu'il est placé seulement dans la position du consommateur à qui l'on demande son avis sur la qualité des prestations reçues. Il y a un réel contresens à inscrire l'éducation dans un tel modèle de marché car l'étudiant ne peut pas être ramené à un client (McMurtry, 1991, Ritzer, 1996, Rowley, 1996). C'est par son engagement actif dans l'enseignement qu'il participe à la qualité de celui-ci. Deux types de résultats significatifs à cet égard sont avérés. D'une part, les cours correspondant aux centres d'intérêt des étudiants sont généralement mieux évalués (Feldman, 1976, Marsh, 1983, Marsh et Dunkin, 1997). C'est même le tiers de la variance des résultats des EEE qui est expliqué par la variable « intérêt pour la matière » ; cette variable, en conduisant les étudiants à développer une motivation accrue pour le cours, est véritablement constitutive de la qualité de l'enseignement. D'autre part, et contrairement aux représentations courantes, ce sont les cours les plus exigeants parce qu'ils correspondent à des attentes élevées des enseignants et à une charge importante de travail des étudiants, qui recueillent les meilleures appréciations (Brady, 1994). Ces différents éléments manifestent que l'enseignement est le fruit d'une coproduction entre l'enseignant et les étudiants et qu'il ne peut être ramené à la seule prestation de l'enseignant. Une des conditions de la réussite de l'enseignement repose sur ce partenariat qui nécessite une claire vision par chacun de ses responsabilités respectives. Une situation dans laquelle les étudiants n'auraient que des droits irait à l'encontre d'une réalité évidente, à savoir la co-responsabilité de l'étudiant quant à son apprentissage. Cette critique est bien sûr dépendante de la conception de l'EEE. Si elle prétend au statut de « jugement étudiant » sur l'ensemble du processus enseignement-apprentissage (y compris ses résultats qui dépendent étroitement de l'investissement de l'étudiant), elle est fondée. Si l'on veille clairement à ce que l'EEE se cantonne à une prise d'informations auprès des étudiants sur des aspects du processus d'enseignement qui sont sous la responsabilité de l'enseignant (clarification du contrat, méthodologie variée...), la critique perd de sa pertinence.

3.1.2. L'enjeu d'une articulation du quantitatif et du qualitatif pour l'appréhension des conditions de milieu et de sens

La volonté institutionnelle d'obtenir des indicateurs quantitatifs ainsi que la tentation de contrôler le système de formation se sont progressivement imposées. Le dispositif informatisé d'EEE a été jugé utile par tous les responsables de l'université qui l'ont pratiqué et les étudiants en ont été

plutôt satisfaits. Il a néanmoins aussi focalisé les oppositions et les résistances des enseignants, étant donné qu'il a accentué la visibilité et la publicité des points de vue des étudiants. D'autre part, certaines composantes de l'université ne se sont pas engagées dans ce dispositif, optant pour leur propre système afin peut être de pouvoir contrôler les résultats fournis à la présidence. De plus, l'absence de culture d'une évaluation formative, aussi bien de la part des responsables, des enseignants que des étudiants, a pu contribuer à une dérive progressive vers une bureaucratisation du processus, à savoir vers un dispositif plus formel que réellement créateur d'une dynamique collective autour de la qualité de la formation. La culture pédagogique dominante, individuelle et disciplinaire s'est révélée antagoniste à des pratiques interdisciplinaires concertées dont certains enseignants étaient demandeurs, mais qui sont restées minoritaires. Car le passage à l'évaluation formative de l'enseignement ne va pas de soi. Il nécessite de forts changements et notamment d'inclure le point de vue des étudiants dans une dynamique globale d'optimisation. Les informations recueillies auprès des étudiants ont suscité des réactions contrastées, aussi bien en raison de leur caractère non univoque que de leur dimension quelquefois critique, voire très critique, vis-à-vis de l'enseignement reçu. Dans de nombreux cas, l'information quantifiée est restée floue et le faible nombre de remarques exprimées dans les questions ouvertes n'a pas permis de comprendre réellement ce que voulaient dire les étudiants. En outre, les appréciations ont pu quelquefois apparaître contradictoires et peu utiles en tant que telles pour l'amélioration de l'enseignement.

Étant données les importantes variations d'appréciation entre les étudiants et l'anonymat des questionnaires, la signification des moyennes et des profils de réponses est restée souvent obscure et difficile à interpréter sans information supplémentaire. Ce sont des séances organisées à propos de l'EEE entre des enseignants volontaires et des étudiants qui ont permis de préciser le sens donné aux questions et aux réponses. Ces séances qualitatives, bien que trop rares, complémentaires aux questionnaires anonymes se sont révélées particulièrement fécondes, étant donné qu'elles ont permis une construction collective du sens même de l'enseignement.

L'analyse de cette médiation sociotechnique située a conduit à mettre l'accent sur les enjeux de l'articulation de données quantitatives à des données qualitatives contextualisées. L'enquête standardisée basée sur des questionnaires portant sur les attitudes comparées des enseignants, des étudiants et des responsables ainsi que sur les effets perçus par ces mêmes acteurs a été significativement complétée par des analyses d'entretiens ciblés et des incidents critiques repérés au cours de l'observation participante, qui ont permis d'apprécier à la fois l'impact de la culture du milieu en matière d'EEE ainsi que la dynamique interactive du processus et du sens donné à l'évaluation.

Mais la prise en compte de la subjectivité et des conditions du milieu nécessitent d'autres formes de dispositifs de médiations élargis aux données de subjectivation. Il est apparu que les entretiens ont été particulièrement appropriés pour révéler les ambiguïtés et les paradoxes de l'EEE ainsi que l'ambivalence des positionnements, allant de l'intérêt affirmé au rejet manifeste. Par ailleurs, quelques expériences menées avec des enseignants volontaires, et sous leur contrôle, se sont révélées fructueuses et satisfaisantes pour les enseignants comme pour les étudiants, car elles ont permis d'approfondir les points de vue singuliers et l'impact du contexte.

La recherche a mis en évidence la diversité des représentations et comportements des enseignants et leur importance dans la compréhension du dispositif d'évaluation, et de son évolution. Les attitudes des enseignants envers l'EEE étaient variables et fortement corrélées à leurs positionnements idéologiques, quant à la place des étudiants dans l'enseignement. Les enseignants qui exprimaient des attitudes négatives au sujet de l'EEE ont aussi été ceux qui considéraient que les étudiants ne pouvaient pas contribuer avec l'enseignant à la construction de son enseignement. D'autres qui estimaient au contraire que les étudiants pouvaient y contribuer étaient favorables à l'EEE.

Toutefois, l'impact des facteurs individuels dans les évaluations des étudiants et l'importance de la réflexion quant aux aspects identitaires, mis en jeu dans des dispositifs d'évaluation, n'ont pas été explorés alors que ces dimensions auraient pu permettre de mieux appréhender la compréhension des phénomènes et favoriser ainsi le développement d'une évaluation écologique. De tels approfondissements sur les conditions de milieu et d'engagement nous semblent déterminants.

3.2. Evaluation interactive avec les boitiers de vote électroniques

La recherche dont il est question dans cette partie cible les pratiques d'évaluation des enseignants universitaires à travers l'étude de leur usage d'un outil d'évaluation interactive, les boitiers de vote électronique. A l'heure où les universités misent fortement sur le numérique et sur l'évolution des pratiques enseignantes vers la pédagogie active, le besoin de recherches dans ce domaine est patent pour « améliorer l'intelligibilité des situations de formation instrumentée et fournir des outils conceptuels utiles à la décision et à l'action » (Albéro, 2014, p 29). Il s'agit donc d'identifier et de comprendre comment les enseignants s'approprient une technologie numérique et les conditions de l'amélioration de cette utilisation.

Une première étape a consisté à répertorier les usages possibles des boitiers dans les cours universitaires (Detroz et Younès, 2014), la suivante s'est concentrée sur la mise à jour des liens existant entre l'appropriation de cette technologie et le niveau de développement techno

pédagogique des enseignants dans le cadre d'un dispositif de recherche-accompagnement (Younès, Soulier et Detroz, 2017).

3.2.1. Proposition d'un cadre intégrateur

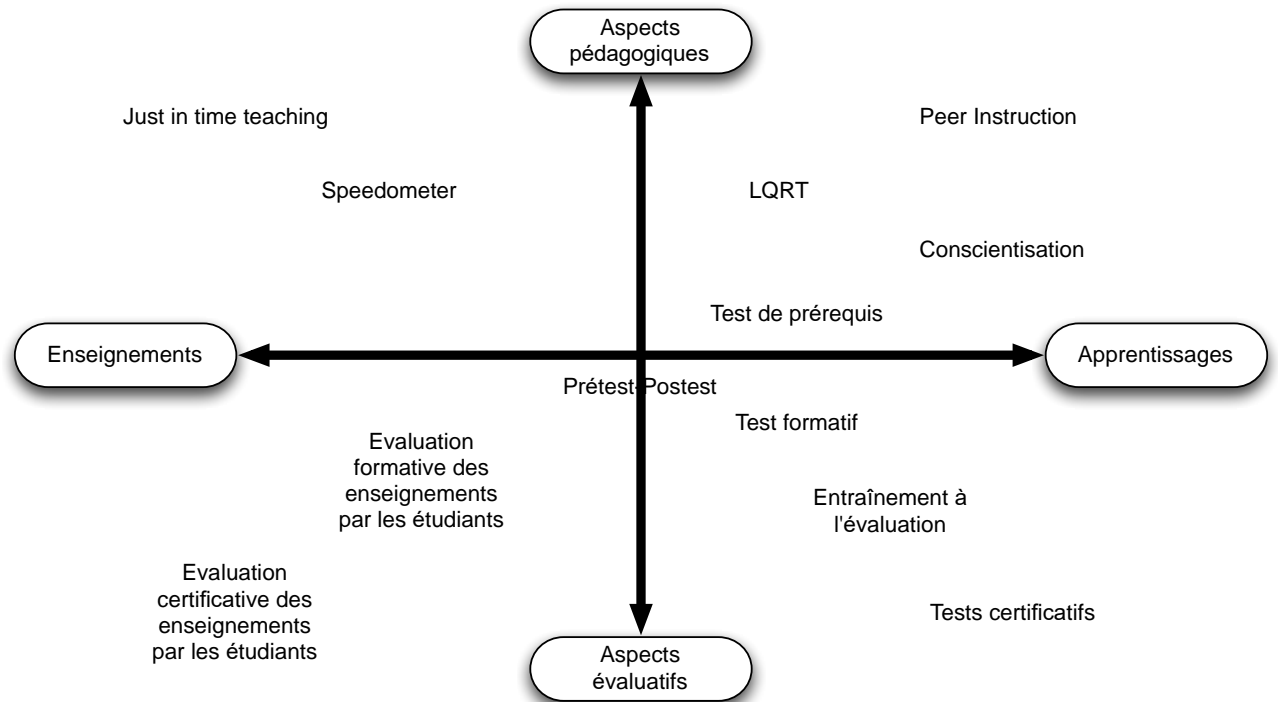
L'usage des boitiers de vote, qui se généralise dans les cours universitaires (Abrahamson, 2006), a donné lieu à de nombreuses recherches qualitatives et/ou quantitatives (pour une présentation en français, voir Dionne, 2012). Différents types de résultats ont été mis en évidence :

- les BVE élèvent le niveau d'intérêt des étudiants, rendent le cours plus interactif, et suscitent un niveau de participation accru ,
- les étudiants apprécient leur utilisation durant les cours,
- l'usage des boitiers entraîne généralement une amélioration des résultats académiques même si certaines études ne retrouvent pas ce résultat

Cette technologie semble donc apporter une valeur ajoutée à la méthode du questionnement, probablement du fait de la rétroaction immédiate qu'elle permet, notamment dans le cadre de l'évaluation formative. Elle offre aux étudiants et à l'enseignant la possibilité de vérifier les connaissances des étudiants et leur compréhension tout au long du cours, ce qui permet la régulation des enseignements et des apprentissages (Brewer, 2004 ; Trees et Jackson, 2007). Les BVE permettent de recueillir une information collective et permettent à l'étudiant de se comparer à ses pairs, tout en préservant son anonymat (DeBourgh, 2007 ; Moredich et Moore, 2007) et donc sans avoir peur du ridicule (Draper et Brown, 2004).

Notons toutefois que les contextes d'utilisation et les usages réellement mis en œuvre sont très peu décrits dans l'ensemble de ces travaux. Pourtant, il semble que, au final, ce que les étudiants apprennent grâce aux boitiers de vote soit largement dépendant de la pédagogie de l'enseignant (Crossgrove et Curan, 2008 ; Dangel et Wang, 2008 ; MacGeorge et al., 2008 ; Tree et Jackson, 2007 ; Suchman et al., 2006).

Nous avons alors proposé un cadre intégrateur des différentes utilisations des BVE défini à partir de la littérature et des observations conduites dans deux universités européennes selon deux axes définissant 4 cadrans (Detroz et Younès, 2014). Le premier cadran est celui où les BVE soutiennent des démarches pédagogiques impactant les enseignements. Dans le second, les BVE sont utilisés pour fournir de l'information aux étudiants en lien avec les aspects pédagogiques de leurs apprentissages. Dans le troisième, les BVE sont en lien avec l'évaluation des enseignements. Enfin, le quatrième, vise à évaluer les apprentissages des étudiants.



Nous avons ensuite étudié de manière qualitative comment les enseignants s'approprièrent effectivement ces usages, en convoquant en particulier les modèles de développement professionnel technopédagogique.

3.2.2. Développement professionnel techno pédagogique et usage des BVE

Des chercheurs s'intéressant à l'usage du numérique dans le domaine éducatif se sont également penchés sur les étapes de développement des enseignants qui découvrent, puis s'approprient les technologies (Sandholtz et al., 1997; Depover et Strebelle, 1997; Moersch 2001; Morais, 2001 ; Raby, 2004)¹⁰. Ces différents modèles proposent une vision essentiellement techno centrée. En parallèle, de nombreuses recherches ont concerné l'impact de l'utilisation des technologies sur l'apprentissage des étudiants. Leurs résultats ne sont pas univoques. Comme le souligne Lebrun

¹⁰ Il existe de nombreux modèles théoriques ou empiriques tentant d'éclairer le développement professionnel des enseignants. Historiquement, les premiers auteurs envisagent une perspective développementale, souvent sous forme de stades (Fuller, 1969 ; Katz, 1972 ; Hall et al., 1979 ; Huberman, 1989 ; Fessler, 1992 ; Nault, 1999). Indépendamment, d'autres se sont intéressés à la professionnalisation des enseignants en ce centrant sur leurs compétences et les dispositifs de formation favorables à leur développement. D'abord essentiellement focalisées sur les savoirs cognitifs et procéduraux, les études se sont progressivement élargies aux compétences métacognitives et psychosociales nécessaires à la fonction d'enseignant. Ainsi, celui-ci n'est plus seulement celui qui doit connaître ou savoir-faire, il est celui qui doit se penser dans l'action et par l'action et qui, par un regard réflexif, est un acteur capable de cerner et maîtriser la complexité du monde qui l'entoure pour que ses actions contribuent au mieux à l'apprentissage de ses étudiants. Cette posture, très proche de la recherche-action, trouvera le terreau à travers lequel prendra racine le Scholarship Of Teaching and Learning dans l'enseignement supérieur. Certains auteurs ont tenté de théoriser ce développement professionnel. Ainsi, après Paquay (1994), Donnay (2002, 2006) propose un continuum du praticien au chercheur en passant par le praticien réflexif (Jorro aussi).

(2011), il ne suffit pas de mettre une couche technologique sur l'enseignement habituel pour créer un amalgame salubre. Il soutient qu'une pédagogie adaptée et rigoureuse est nécessaire pour que les potentialités des outils technologiques soient mobilisées à bon escient et ainsi profitent directement aux étudiants. Plusieurs enquêtes réalisées par Albero et ses collaborateurs (Albero, 2014, p. 29) dans les années 2000 ont d'ailleurs montré que si l'intégration des technologies permettait d'optimiser les pratiques d'enseignement, elle ne conduisait pas pour autant nécessairement à leur transformation ; de la transmission des connaissances au développement des capacités à apprendre.

Le modèle SAMR (Puentedura, 2009) prend quelque distance par rapport aux modèles publiés jusqu'alors, en ce qu'il est résolument techno pédagogique. Il se présente sous la forme d'une échelle à 4 niveaux caractérisant les niveaux d'intégration techno pédagogique :

- La Substitution (la technologie ne fait que se substituer à l'existant)
- L'Augmentation (la technologie entraîne également des améliorations fonctionnelles).
- La Modification (amélioration d'un certain nombre d'activités pédagogiques)
- La Redéfinition (la technologie concernée permet de concevoir des refontes pédagogiques impossibles à envisager sans son intégration).

Karsenti (2014) propose un modèle analogue (ASPID). Il est plus complet en ce qu'il décrit une phase possible supplémentaire : la détérioration (l'usage entraîne plus d'inconvénients qu'il n'apporte de solutions).

Contrairement aux autres, le modèle TPaCK (Mishra et Koehler, 2006) dans le prolongement des travaux de Shulman (1986) ne présente pas ses éléments sous forme d'étapes, mais il offre un cadre conceptuel pour penser le lien entre contenu, pédagogie et technologies de l'éducation. Il a, de plus, l'avantage d'avoir été pensé dans le domaine de l'enseignement supérieur. Selon ces auteurs, l'expertise en enseignement dépend de l'accès rapide à trois systèmes de connaissances très organisés qui concernent [1] le contenu à enseigner (Content), [2] la pédagogie (Pedagogy), [3] les technologies (Technology). Shulman avait nommé l'intersection entre [1] et [2] « Pedagogical Content Knowledge (PCK) ». En France, nous pourrions parler de didactique. Comme le montre la figure 1, trois autres intersections existent : Technological Content Knowledge (TCK), Technological Pedagogical Knowledge (TPK) et Technological Pedagogical Content Knowledge (TPCK).

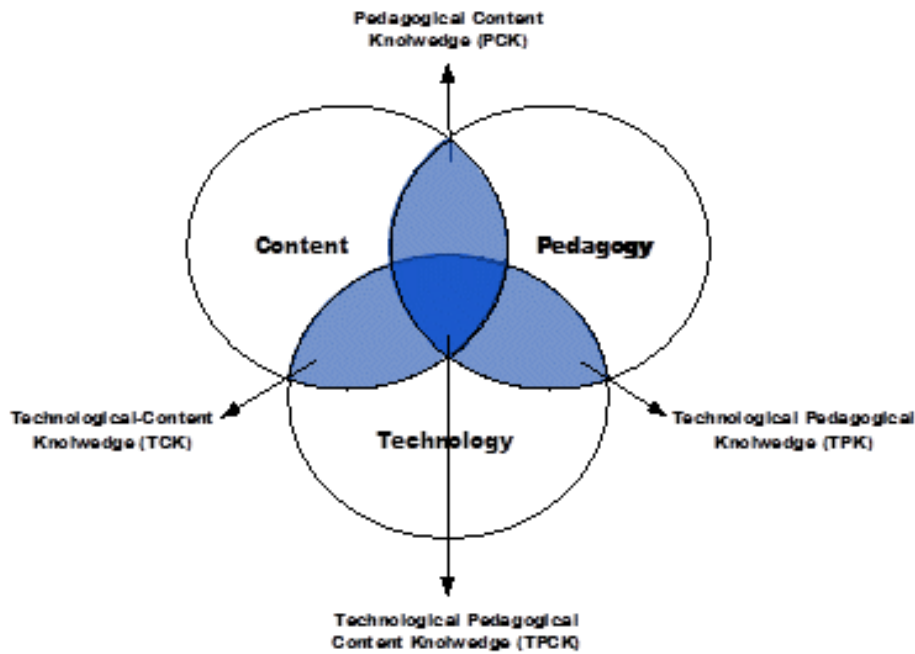


Figure 1 : The TPACK model (Mishra et Koehler, 2006)

Nous avons intégré les perspectives du développement professionnel et du développement technopédagogique et pris en compte l'influence du contexte de l'enseignement, généralement peu considéré dans les modèles de développement professionnel recensés, à travers 9 études de cas d'enseignants universitaires volontaires de différentes disciplines et différents contextes de formation. Le protocole a été élaboré afin d'appréhender finement les dynamiques d'évolution des usages des BVE en lien avec le développement professionnel des enseignants et leur contexte professionnel.

Les résultats montrent sans surprise que le niveau de développement professionnel de l'enseignant oriente son usage des BVE qui se révèle plus ou moins orienté vers le soutien de l'apprentissage des étudiants. Les enseignants préoccupés d'abord par leur façon d'enseigner utilisent les BVE surtout pour de l'évaluation sommative ou de l'entraînement à l'évaluation, mais pas à des fins de régulation de l'enseignement ou de l'apprentissage. En revanche, les enseignants centrés sur l'apprentissage des étudiants, utilisent les BVE dans le but de l'améliorer via des modalités d'évaluation formative, de conscientisation des étudiants et d'ajustement du cours en fonction de leurs réponses. Ce résultat précise des recherches anciennes qui montrent que la centration sur l'étudiant correspond à un stade avancé du développement professionnel (Fuller 1969, Katz 1972, Hall et al., 1979, Huberman 1989, Fessler et Christensen, 1992 et Nault 1999, a et b) et plaide en faveur d'une réelle intégration de la technologie et de la pédagogie. Il permet

également de nuancer les résultats rapportés dans la littérature, qui ont tendance à mettre l'accent sur les apports pédagogiques des BVE car nous montrons que ces effets potentiels sont fortement dépendants de la représentation que l'enseignant se fait du processus d'enseignement apprentissage et de la place de l'évaluation dans ce processus. Nous avons pu répertorier 9 catégories d'impact déclaré de l'usage des boîtiers dans les discours des enseignants aussi bien sur l'enseignant (réflexion pédagogique, identification du niveau de compréhension des étudiants et régulation de l'enseignement) que sur les étudiants (participation et attention accrues, autoévaluation rendue possible, travail personnel augmenté et mémorisation accrue). Ces observations conformes aux résultats rapportés dans la littérature sur l'effet des BVE. Mais ces différentes catégories d'impact n'ont pas été perçues par tous les enseignants de l'étude, elles sont dépendantes de l'usage réel des boîtiers, lui-même dépendant de ce que nous avons identifié comme un niveau de développement professionnel.

3.2.3. Dispositif de recherche/accompagnement de l'usage des BVE et développement professionnel pédagogique

À partir du moment où les technologies doivent être en interactions directes avec les savoirs liés au contenu et avec les aspects pédagogiques pour donner leur pleine mesure, cela signifie que la porte d'entrée technologique permet, lorsque l'accompagnement est performant, un développement professionnel de l'enseignant sur l'ensemble des facettes du modèle, ou du moins sur leurs intersections. Cette thèse est défendue par Puentedura (2014) pour qui le niveau « Modification » peut être atteint en augmentant le niveau des savoirs liés au *Pedagogical Content Knowledge* (PCK), au *Technological Content Knowledge* (TCK), et au *Technological Pedagogical Knowledge* (TPK) alors que pour atteindre le niveau « Redéfinition », il est nécessaire d'entamer une réflexion au niveau du *Technological Pedagogical Content Knowledge* (TPCK).

Dans cette perspective, le développement techno-pédagogique des enseignants et donc l'introduction des technologies, dans leur enseignement, devient un acte de développement professionnel ayant potentiellement de nombreuses répercussions positives sur leur expertise en tant qu'enseignant. C'est selon ce principe que nous avons accompagné plusieurs enseignants dans l'utilisation des boîtiers de vote électronique.

Nous avons intégré les perspectives du développement professionnel et du développement technopédagogique et pris en compte l'influence du contexte de l'enseignement, généralement peu considéré dans les modèles de développement professionnel recensés, à travers l'étude des trois hypothèses de cette recherche. Nous nous attendions tout d'abord à ce que les enseignants s'approprient différemment la technologie en fonction de leur stade de développement professionnel, les enseignants positionnés à un stade de développement professionnel

pédagogique avancé devraient favoriser une utilisation des BVE plus centrée sur l'apprentissage des étudiants que les enseignants moins avancés. Nous nous attendions également à ce qu'une activité réflexive soutenue par la médiation sociotechnique activée par un dispositif d'accompagnement, entraîne un développement professionnel dans le sens d'une prise en compte accrue du processus d'apprentissage dans l'utilisation des BVE. Enfin nous avons étudié en quoi les caractéristiques du contexte d'enseignement, selon que celui-ci soit plus ou moins orienté vers le soutien à la pédagogie, modulent les effets du dispositif d'accompagnement.

Ces hypothèses ont été étudiées à travers une recherche qualitative reposant sur 9 études de cas d'enseignants universitaires volontaires de différentes disciplines et différents contextes de formation. Le protocole a été élaboré afin d'appréhender finement les dynamiques d'évolution des usages des BVE en lien avec le développement professionnel des enseignants et leur contexte professionnel selon des modalités d'accompagnement léger (4 enseignants) versus approfondi (5 enseignants). Dans le cas de l'accompagnement dit léger, les enseignants ont été interviewés à deux reprises et observés en cours. L'accompagnement approfondi a visé à favoriser une réflexion soutenue dans le temps à partir de trois sources d'information : des entretiens réitérés y compris d'autoconfrontation au film d'une séance de cours, l'observation de l'enseignement et son enregistrement ; le point de vue des étudiants (voir Figure 2).

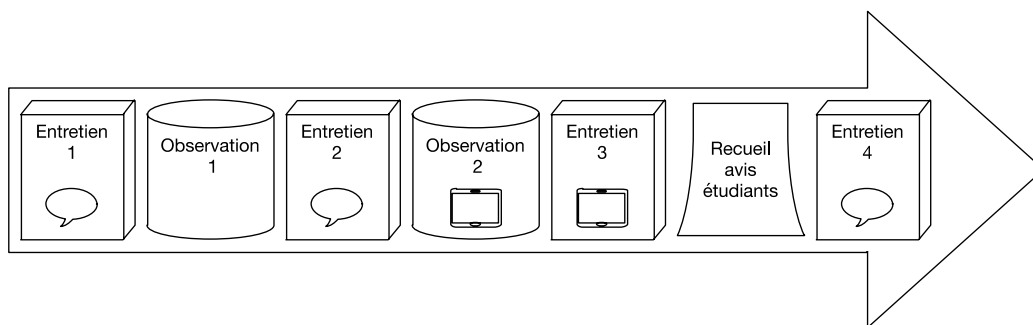


Figure 2 : Protocole d'accompagnement approfondi

Mais l'objectif premier de cette recherche était d'identifier les effets d'un dispositif d'accompagnement sur le développement professionnel pédagogique de l'enseignant. Rappelons que nous avons distingué deux modalités d'accompagnement (léger vs approfondi). La présentation détaillée de l'outil peut, en soi, être assimilée à une forme d'accompagnement dans la mesure où cette présentation révèle le potentiel de l'outil et son intérêt pédagogique. De plus, le questionnement de l'enseignant sur ses choix d'utilisation des BVE et sur les effets de ces choix peut être considéré comme la première étape d'un support d'évolution professionnelle. Le dispositif d'accompagnement approfondi a conduit à des évolutions pédagogiques plus importantes. Ainsi, dans le cadre de ce dispositif, nous avons accompagné l'enseignant dans le déploiement d'une technologie mais lui avons également offert une batterie d'outils réflexifs. Or,

comme le rapporte Day (1999), le développement professionnel est rendu possible par les expériences d'apprentissage spontanées de même que par celles conscientes et planifiées. En permettant l'établissement de liens entre l'intention initiale de l'enseignant, son action perçue et réelle et l'impact de son action, cet accompagnement offre de réelles opportunités de développement.

Il apparaît distinctement que les enseignants ayant bénéficié d'un accompagnement approfondi perçoivent davantage le caractère pluriel de l'impact de l'usage des BVE sur leur enseignement que ceux ayant bénéficié d'un accompagnement léger.

De plus, à l'issue de l'expérimentation, tous les enseignants ont fait évoluer leur utilisation des BVE mais de manière plus ou moins importante. Les enseignants bénéficiant de l'accompagnement léger ont enrichi leur utilisation en associant une modalité d'évaluation formative à leur objectif certificatif initial. Cette évolution est encore plus manifeste chez les enseignants inclus dans le protocole d'accompagnement approfondi qui les a conduit à développer leur usage des BVE à plusieurs niveaux :

- En amont de l'usage en intégrant les BVE à une réflexion sur la structuration de leur enseignement ou en évoluant dans la finalité des questions posées. On voit ainsi un enseignant évoluer de la conception de question à réponse unique vers la conception de questions conçues pour initier un débat
- Durant l'utilisation en cours en prenant conscience de la nécessité de fournir un feedback immédiat aux étudiants et en réfléchissant sur les modalités que celui-ci peut prendre pour faciliter les apprentissages.
- En aval dans l'ajustement du cours relatif aux réponses des étudiants.

Au-delà de ces améliorations, le protocole d'accompagnement approfondi a conduit les enseignants à se questionner plus généralement sur leur pédagogie suscitant une prise de conscience propice à un meilleur alignement de l'enseignement entre objectifs, méthodes pédagogiques et évaluation des apprentissages. Ils constatent par exemple que certains points essentiels du cours n'ont pas été abordés, ou ils découvrent des difficultés inattendues ou encore que les objectifs de l'enseignant ne sont pas partagés par les étudiants.

La progression globale des niveaux d'intégration techno-pédagogique des BVE, parfois jusqu'au stade le plus élevé (Redéfinition) - et, si l'on se réfère au modèle TPaCK, aboutissant à renforcer le *Technological Pedagogical Content Knowledge* - atteste de l'efficacité de l'accompagnement approfondi et de son impact positif sur l'ensemble du développement professionnel. On peut donc considérer, à ce titre et au regard des résultats obtenus, que la porte d'entrée technologique

permet le développement professionnel de l'enseignant sur l'ensemble des facettes du modèle TPaCK aux conditions que l'accompagnement soit performant et qu'il soit maintenu dans la durée. Reste que l'utilisation accompagnée d'un outil a ses limites, quelle que soit la qualité de l'accompagnement. Le poids des croyances et de la culture professionnelle a conditionné, dans notre étude, la perception du potentiel pédagogique de l'outil. Enfin, il apparaît clairement que les contextes social, administratif et technique d'un enseignant sont fortement liés à la réussite de l'appropriation durable de l'outil. Les ressources organisationnelles et matérielles sont également cruciales ainsi que l'existence d'une dynamique pédagogique collective.

Si la durée de l'expérience exige de rester prudents à l'égard des conclusions présentées, nous pensons que l'utilisation accompagnée des boitiers est susceptible d'améliorer la pratique d'enseignement sur le moyen et le long terme. Des enseignants réalisent qu'en modifiant leur pédagogie (en optant pour la classe inversée par exemple), ils atteignent mieux les objectifs visés relativement à une approche classique de transmission du contenu par l'enseignant. Dans ce contexte, le questionnement et les échanges générés en cours n'entraînent pas la « perte de temps » redoutée et souvent invoquée comme obstacle à la mise en place de méthodes actives, puisque l'enseignant cible les points à développer et passe sur ce qui a été acquis, l'évaluation avec les BVE lui permettant d'ajuster son cours. On passe d'un enseignement davantage basé sur « je transmets des connaissances » à un enseignement qui se préoccupe en continu de la compréhension et des connaissances des étudiants. C'est le passage du *teaching* au *learning* d'une certaine manière (Biggs, 1991). Reste que cette prise de conscience, si elle est quasi-systématique, n'est pas toujours suffisante pour inciter l'enseignant à réfléchir sur l'amélioration de son dispositif pédagogique et, in fine, à envisager sa modification ou sa redéfinition. C'est ici que l'accompagnement approfondi sur le long terme jouerait un rôle primordial.

Constatons que si de nombreuses recherches scientifiques ont vu le jour sur l'intégration et l'impact des nouveaux outils numériques dans l'enseignement universitaire, assez peu proposaient d'observer les effets d'une utilisation accompagnée de ces outils sur l'enseignement et le développement professionnel des enseignants. L'importance du contexte professionnel dans la réussite de l'intégration et l'utilisation de ces outils au quotidien a également été assez peu abordée. L'objectif de la recherche dans ce domaine a donc été de répondre à ces interrogations. L'expérience conduite a permis de conclure à un premier constat qui, s'il paraît évident, n'en demeure pas moins une première donnée essentielle : l'utilisation accompagnée d'un outil facilite son intégration dans l'environnement pédagogique de l'enseignant. Il participe aussi à améliorer sa pédagogie. Par améliorer, entendons une évolution très nette de la façon d'utiliser et d'envisager les BVE, d'une pratique d'évaluation exclusivement certificative à des pratiques

d'évaluation formative et de conscientisation tournées vers la prise en compte de l'apprentissage des étudiants. C'est la découverte renseignée et l'apprentissage techno-pédagogique d'un nouvel outil, associés à une réflexion sur sa façon d'enseigner qui, à terme, participent au développement professionnel de l'enseignant.

Reste que la réussite de l'intégration techno-pédagogique de l'outil et ses effets sur le développement professionnel sont soumis à deux conditions. L'accompagnement, associant dimension technique et dimension pédagogique, requiert d'être soutenu de manière pérenne. Par ailleurs, le contexte professionnel (pédagogique, administratif, technique) constitue une autre condition déterminante. On découvre en effet rapidement les limites du choix des technologies et du développement professionnel comme leviers d'adaptation quand ils sont confrontés à des contraintes qui viennent empêcher l'efficacité et occupent une place prégnante dans les choix pédagogiques des enseignants.

Comme le souligne Annoot (2012, 2014), les usages du numérique dans les pratiques d'enseignement à l'université interrogent le système universitaire dans ses différentes dimensions, politiques, sociales et pédagogiques. Dans la recherche sur la conception d'un dispositif d'EEE informatisé ou dans celle portant sur l'usage d'un outil d'évaluation interactive, l'étude des interactions et correspondances entre ces différentes dimensions requiert des enquêtes situées, mobilisant différentes méthodes, qualitatives et quantitatives, pour tenter d'appréhender les conditions de milieu et de sens qui définissent les potentialités transformatives de ces dispositifs d'évaluation, émancipatrices ou aliénantes.

Partie 4 - Dispositifs d'évaluation collaboratifs : transformations individuelles et collectives

La perspective de régulation critique de l'action (Hadji, 1992, De Ketele, 2001), est déterminante pour s'adapter à un contexte toujours changeant. Ainsi que le développe Hartmut Rosa dans son ouvrage récent *Résonance, une sociologie de la relation au monde* (2018), la société capitaliste moderne, caractérisée par des processus d'accélération et d'accroissement généraux met en crise les multiples formes de résonances : « La logique d'accroissement contamine à ce point la conduite de vie des sujets qu'ils sont contraints d'aller toujours plus vite, ne serait-ce que pour maintenir leur place dans le monde » (p31). Concurrence, compétition, critères de performance, quantification, entraînent une course effrénée imprégnant toutes les situations professionnelles et personnelles au détriment d'un équilibre. Avec le thème de la résonance, Hartmut Rosa prospecte des manières alternatives de vivre et de faire, considérant que « les rapports de résonance présupposent continuellement un ajustement rythmique réciproque et doivent satisfaire à des exigences spécifiques de synchronisation. » Définissant la résonance comme « se sentir en harmonie » (p.75), il explore l'art d'entrer en relation de manière cognitive et affective, réintroduisant l'importance des rapports interpersonnels et de sens : « En tant qu'elle dénote une faculté de sympathie et d'empathie, la résonance crée et signale une demande d'interaction et de coopération, donc de capital social » (p.39).

Faire face aux transformations de l'enseignement supérieur (Romainville, 2004, 2014) suppose une adaptation aussi bien individuelle que collective. Elle se nourrit alors de différentes formes de partage et de mise en commun du sens entre les différents acteurs de l'université (étudiants, enseignants, responsables, accompagnateurs pédagogiques...). On peut constater que malgré une culture du coopératif qui prend de l'ampleur, son déficit aux différents niveaux d'enseignement reste un chantier d'envergure (Marcel, 2009). En effet, les implications et les interactions des uns avec les autres s'avèrent incontournables dans une évaluation conçue en lien avec des dynamiques de transformations individuelles et collectives. Un projet d'évaluation de l'enseignement appropriable repose sur la nécessité d'une concertation permettant une prise en compte des acquis de l'expérience de chacune des personnes impliquées. Les dispositifs d'évaluation comme médiations psycho-socio-techniques se trouvent alors confrontés au tournant du partage dans la façon de penser l'évaluation de l'enseignement, du point de vue aussi bien théorique, pratique qu'opératoire.

Le mot de partage est en lui-même théorique, politique et méthodologique en ce qu'il désigne l'action de diviser quelque chose en plusieurs parties mais aussi l'action de mettre en commun, « l'action conjointe », la « participation » (Rey, 1994, p. 1436). Le philosophe Rancière associe ce double sens dans sa pensée du processus démocratique. Dans son ouvrage éponyme de philosophie politique, il précise qu'« un partage du sensible fixe donc en même temps un commun partagé et des parts exclusives » (Rancière, 2000, p.12) requérant la constitution de « scènes polémiques » (Rancière, 1995, p.66) comme espaces de paroles contradictoires où chacun participe de manière égale, y compris celui qui est minoritaire. Il précise d'autre part « comment dans un espace donné, on organise la perception de son monde, on relie une expérience sensible à des modes d'interprétation intelligibles » (Rancière, dans Zerbib, Lévy et Rennes, 2007). Ainsi que le souligne un autre philosophe contemporain, « le sens est dans le partage d'un 'en commun', le partage articulé d'un sens qui donne lieu précisément à la communauté » (Nancy, 1999, p.211). Ce qui fait écho au projet d'« agora du sens » pensé par Arendt (1978, p. 237-238) comme la mise en place de la pluralité dans un espace d'exposition réciproque.

4.1. Le tournant du partage dans l'EEE

Le partage constitue un chantier expérimental aux différents niveaux interindividuel, institutionnel et interculturel. Et c'est sur cet enjeu que nous avons mis l'accent (Younès, 2015), en se centrant sur son importance pour mieux articuler transcontextuel et contextuel ainsi que les logiques institutionnelles, collectives et singulières.

Ainsi, dans la démarche d'EEE, la faculté de jugement de l'étudiant se trouve plus ou moins mobilisée. Par jugement, il est entendu la propriété d'exercer son entendement en situation et de stimuler la faculté psychique intellectuelle qui permet de saisir les problèmes et les situations et de les apprécier. Et c'est l'ensemble de ce processus expérientiel qui participe à une évaluation qui construit des valeurs et qui par là-même devient motrice (Dewey, 1993). Mais l'évaluation de l'enseignement est aussi une démarche d'ensemble qui amène à se démarquer d'une vision de l'EEE se limitant à une collection d'appréciations individuelles, qui estompe non seulement les points de vue singuliers et les écarts en les ramenant à une simple moyenne, mais aussi les dynamiques de mise en commun. Les conditions d'engagement des enseignants et des étudiants dans un processus de dialogue et de formalisation du processus d'évaluation sont donc déterminantes pour sa fécondité. Réduite à une procédure formelle et stéréotypée de traitement d'appréciations individuelles sur la base d'un référentiel normatif, l'évaluation de l'enseignement par les étudiants constitue une forme appauvrie de contribution au projet d'amélioration de l'enseignement universitaire, qui passe par la construction d'un en commun pédagogique.

L'analyse de l'abondante littérature sur le sujet et la variété des démarches d'EEE dont nous avons rappelé les principales orientations invite à la repenser entre généralité transcontextuelle et spécificité contextuelle, ainsi qu'entre subjectivation et intersubjectivation. Une orientation féconde consiste donc à activer les dynamiques de partage à tous les niveaux du processus que ce soit en amont dans la définition des objectifs, des critères et des modalités de l'EEE ou en aval dans l'élaboration d'un sens partagé à travers la mise en relation et l'explicitation de la multiplicité des points de vue.

4.1.1. Des figures de partage

Nous présentons les différentes figures de partage que nous avons repérées tout au long du processus d'EEE (en amont, durant son effectuation et en aval, à partir de l'analyse des expériences belges, suisses, québécoises, et françaises.

- ***Partage de la responsabilité dans le processus de mise en œuvre de l'EEE***

Les établissements suisses d'enseignement supérieur disposant d'une forte autonomie pour l'implantation de leurs démarches qualité, ont des pratiques présentant divergences et convergences. La recherche de Dumont, Rochat, Berthiaume et Lanarès (2012) a montré que la Haute école d'ingénierie du canton de Vaud (HEIG-VD/HES-SO) est plus orientée vers le contrôle que l'université de Lausanne. En effet, à la HEIG-VD/HES-SO, l'EEE est pilotée par l'institution, les résultats en sont communiqués aux enseignants mais aussi au responsable hiérarchique et ils sont pris en compte dans la titularisation des enseignants. En revanche, à l'université de Lausanne, les enseignants disposent d'une plus grande responsabilité dans le choix et les modalités de l'EEE, et ils sont les destinataires exclusifs des résultats des EEE. Cependant les deux institutions ont opté pour soutenir le développement professionnel des enseignants en mettant en place des structures dédiées à leur formation, au conseil pédagogique et à la recherche en pédagogie universitaire. Dans ces deux établissements également, les enseignants ont la possibilité de construire des questionnaires « sur-mesure » avec l'aide d'un conseiller pédagogique, avec une marge de manœuvre sur le processus d'EEE. Les chercheurs de ces structures constatent dans ces contextes une augmentation régulière de la pratique volontaire de l'EEE ainsi que des demandes de formation en pédagogie. Ils notent également des effets perçus comme positifs par les étudiants et les enseignants.

Finalement à partir d'un cadre de référence commun, les ajustements localisés et encouragés qui ont lieu au niveau institutionnel comme au niveau individuel contribuent au renforcement d'une culture de l'EEE au service du « développement professionnel des enseignants », et ce quelles que soient les différences dans la manière de procéder des deux établissements.

- ***Partage en amont avec les enseignants de la construction des critères de qualité de l'enseignement***

En Belgique francophone, les travaux de Romainville (2009) permettent de préciser une autre modalité de partage dans la culture formative de l'évaluation. Ainsi décrit-il avec précision l'importance de la construction collective de critères de qualité de l'enseignement comme préalable à l'acceptation et à la pertinence d'une démarche d'EEE, tout en soulignant cependant le caractère partiel et limité dans ses effets. Ce qui conduit à faire l'hypothèse que le renforcement d'une mise en commun pédagogique à d'autres niveaux du processus d'évaluation (notamment l'accès aux résultats et le suivi de l'évaluation) permettrait d'étudier l'impact d'un partage élargi.

- ***Partage en amont avec les étudiants de la construction des critères de qualité***

Abernot, Gangloff-Ziegler et Weisser (2012), rendent compte d'une expérience en France à l'Université de Haute Alsace visant à fonder les questionnaires d'EEE sur les critères de qualité identifiés par les étudiants lors d'entretiens de groupes qui « poussent les individus interrogés à approfondir leurs positions respectives, à construire collégialement leurs critères de qualité ». Ils proposent également de présenter le questionnaire d'EEE aux étudiants en début de semestre afin de s'accorder sur le sens des questions. Se référant à Clot (1999, p.131), ils visent une certaine « validité écologique » des situations étudiées pour « mieux comprendre des phénomènes vécus en commun mais à partir de positions parfois distinctes », de telle sorte que se trouverait favorisés un meilleur ajustement du dispositif d'évaluation aux étudiants et une meilleure appropriation collective.

- ***Partage en aval des points de vue entre étudiants et enseignants***

Dès 1987, Abernot avait proposé un outil permettant de mettre en correspondance l'évaluation de l'enseignement par l'enseignant et par les étudiants en les invitant à situer l'enseignement sur la même grille. Il présente l'outil et la démarche dans Abernot, Gangloff-Ziegler et Weisser (2012). L'évaluation consiste alors plus dans la mise en perspective des différents points de vue que dans une perspective unilatérale, le but étant de susciter une discussion amenant à une meilleure compréhension mutuelle et éventuellement à des adaptations réciproques, et ce à différents moments de l'année universitaire.

- ***Partage en aval de l'analyse et des décisions consécutives à l'EEE pour apprécier et faire évoluer le programme de formation***

La démarche institutionnelle originale mise en place à la Haute École de Namur-Liège-Luxembourg dont rend compte Ernst (2014) consiste à demander à chaque enseignant de fournir aux

directeurs des programmes un « projet d'aménagement » de son enseignement (en termes de contenus, méthodes, modalités d'évaluation...) sur la base d'une analyse incluant notamment l'avis des étudiants. Les directeurs des programmes effectuent des synthèses anonymisées rendues publiques et discutées au sein des conseils pédagogiques. Ainsi, l'EEE, conduite individuellement par chaque enseignant, par sa diffusion filtrée et les discussions afférentes, participe à un processus plus large de régulation collective des formations dispensées aux étudiants. En effet, la comparaison du contenu des « projets d'aménagement » individuels regroupés en fonction des acteurs, des années et des programmes amène à identifier les attentes des étudiants, la place accordée à certaines pratiques d'enseignement-apprentissage, les difficultés communes, les accords et désaccords entre collègues.

Au Québec, Bernard (2011) propose de renforcer le partage autour de l'EEE en la repensant comme un élément à mettre en discussion de telle sorte qu'elle devienne un support d'analyse concertée d'un programme de formation dans sa globalité. Il est insisté sur la nécessité que ces échanges soient basés sur des résultats d'EEE rendus anonymes afin de libérer la parole de chacun et de ménager les personnes.

Notons qu'à propos des figures de partage de l'EEE, en Suisse comme en Belgique francophone, en France ou au Québec, l'accent est principalement porté sur la mobilisation et la réflexivité de l'enseignant individuel, bien que l'on voit émerger aussi l'importance accordée à la part du collectif enseignant et à celle du collectif étudiant dans l'élaboration appropriée de critères comme dans l'analyse consécutive aux EEE. En ce qui concerne la co-construction pédagogique qui pourrait s'ensuivre (Abrami, d'Apollonia et Rosenfield, 2007), elle reste encore peu développée. Ces différentes expériences manifestent également la nécessité d'une méthodologie rigoureuse dans tout le processus d'EEE (Theall et Franklin, 2001 ; Detroz, 2014), en particulier dès qu'il s'agit d'activer le partage.

Il est bien certain que le partage ne fait pas disparaître les tensions inhérentes de l'évaluation, surtout en cas d'évaluation négative. Cela peut entraîner de fortes frustrations mais celles-ci se trouvent exacerbées s'il n'est pas possible d'en débattre, ce qui peut éventuellement entraîner des modifications des manières de faire.

4.1.2. Quelle mise en commun de l'EEE en France ?

Dans notre recherche sur les enseignants-chercheurs (Paivandi et Younès, 2017), deux questions portaient spécifiquement sur la dimension collective de la construction de l'enseignement. S'il est généralement considéré et déploré que l'EEE et plus globalement la pédagogie ne font pas ou peu l'objet d'un travail collectif, on constate cependant de fortes différences d'appréciation quant à l'impact du dispositif d'EEE suivant les déterminations et partages qui le caractérisent.

- *Impact des discussions autour des résultats d'évaluation*

Dans les deux universités de l'enquête, les dispositifs d'évaluation sont différenciés selon les composantes. Dans les composantes où existe un dispositif d'évaluation de l'enseignement structuré associé à une discussion collective des résultats quant à l'évaluation de la formation, les enseignants chercheurs estiment davantage que celui-ci entraîne d'échanges dans la pédagogie. Il s'agit principalement de discussions portant sur l'évaluation de la formation dans sa globalité et non pas sur les EEE bien qu'elles en fassent partie. Cette appréciation d'un impact a été systématiquement retrouvée dans les entretiens d'enseignants-chercheurs appartenant à des composantes de l'université du type école d'ingénieur, qui font également plus état d'une conception de la pédagogie comme démarche collective.

- *Un dispositif institutionnel d'EEE perçu comme n'entraînant pas de dynamique pédagogique collective*

D'une manière générale dans les Unités de Formation et de Recherche (UFR), quelle que soit la filière, le dispositif institutionnel n'est pas jugé comme producteur d'effets positifs en termes de dynamique collective, soit que les évaluations sont conduites sans que les enseignants concernés aient les retours de ces évaluations (dont les résultats arrivent généralement aux responsables des formations), soit que les enseignants reçoivent ces résultats comme des sanctions, soit il s'agit en tout cas d'une affaire individuelle, soit encore les enseignants se sentent menacés par le dispositif mis en place.

- *Une perception différenciée à propos du travail collectif selon la position institutionnelle*

Les enseignants-chercheurs responsables de formation sont plus enclins à considérer qu'une forme de travail collectif autour de la pédagogie et de l'évaluation a lieu. Au niveau des enseignants chercheurs, certains éprouvent une solitude douloureuse même si elle peut être parfois choisie. Les responsables des formations rendent compte de leur souci d'utiliser les résultats pour améliorer la qualité de la formation voire le développement professionnel des enseignants, mais également d'un certain malaise et d'une difficulté patente à cet égard.

4.1.3. Subjectivation et intersubjectivation dans un processus formatif

Ces différentes études montrent la complexité de la notion de partage qui renvoie, comme l'a bien souligné Rancière à la fois à de la séparation entre les choses et à de la mise en commun comme nous l'avons déjà souligné. Cette double face du partage participe du seuil critique de l'EEE en tant que transformation des pratiques pédagogiques et du système d'enseignement. L'inertie et la résistance que l'EEE suscite manifestent l'importance de ce seuil à franchir et

l'ambiguïté des réactions que ce passage provoque (Younès, 2007). Il s'avère particulièrement difficile étant donnée l'extrême complexité engendrée par les différents emboitements de contextes et de valeurs (personnel, groupal, disciplinaire, institutionnel, national, transnational) dans lesquels l'EEE se trouve engagée. Il apparaît que la prise en compte des contextes particuliers permet à la fois de comprendre les résultats contradictoires mis en évidence dans les différentes études et d'adapter des procédures génériques qui sans cela demeurent normatives et peuvent tout simplement rester lettre morte ; au pire se révéler contreproductives voire délétères. Elles ne disent rien si elles ne sont pas réinterprétées par chacun, en situation et selon sa propre histoire. C'est dans cette capacité à adapter en partageant ces processus génériques *in situ* que peut être situé le tournant de l'évaluation. Le but est bien que les enseignants et les étudiants se comprennent et jugent l'enseignement et ses appréciations en les rapportant à un contexte qui fait sens pour eux, mais aussi qu'ils améliorent ensemble à partir d'une compréhension subjectivante et intersubjectivante, l'enseignement et l'apprentissage. Le savoir pédagogique universitaire se construit individuellement et collectivement à partir d'une reconnaissance et d'un débat autour des divergences et convergences institutionnelles, disciplinaires et culturelles ainsi qu'à partir des données d'expérience qui peuvent être confrontées. Cette approche située amène donc à ne pas s'enfermer dans un référentiel conçu comme une « norme » de comparaison prescrite de l'extérieur (Figari, 2001) pouvant être perçue comme particulièrement menaçante et inappropriée mais au contraire à valoriser la diversité des expériences et à faire émerger des orientations concertées.

La mise en mouvement que suscite le partage s'exprime non seulement au niveau de l'enseignant mais aussi à celui de l'étudiant et à celui du milieu de formation, ces trois niveaux étant étroitement liés (Younès et Paivandi, 2017). L'enseignant est mis face à un miroir qui lui renvoie comment son cours est perçu par les étudiants et peut ainsi réaliser les décalages existant entre ses intentions pédagogiques et leur impact. L'étudiant est encouragé à réfléchir aux buts de l'enseignement et à son expérience d'apprentissage en comprenant mieux ce qui la facilite et ce qui l'entrave. Mais les relations enseignant/étudiant s'inscrivent aussi toujours dans un milieu complexe et évolutif qui détermine largement leurs conditions de possibilité. En effet l'enseignement dépend certes de l'enseignant mais également d'une coproduction impliquant tous les acteurs : étudiants, personnels administratifs et décideurs. Dans une telle approche écologique du partage, le milieu de formation se trouve donc ramené au cœur du processus pédagogique. Ces mises en relation contribuent à la création d'un espace collaboratif de reconnaissance et de coopération propre à favoriser une forme de subjectivation grâce à ce que nous pourrions appeler des ateliers d'intersubjectivation durant lesquels les participants prennent

en considération la pensée de l'autre dans un espace de dialogue critique et ainsi s'envisagent à l'autre.

Ce tournant du partage, qui amène à faire tenir ensemble différents types d'environnements administratifs, théoriques et pratiques s'avère bien indissociable d'une évaluation de l'enseignement à visée formative et transformative. En fait, c'est l'évaluation formative qui peut contribuer à un dialogue critique entre les acteurs permettant d'engendrer une intersubjectivité entre les acteurs ayant des positions assez différentes. Un tel type d'évaluation ouvre des espaces de médiation permettant aux différents acteurs du collectif pédagogique de se rencontrer et d'échanger autour de leurs représentations, de leurs pratiques et de leurs évolutions. Ce partage peut contribuer à un éclairage et un soutien de la relation enseignant/étudiant, mais également à nourrir le milieu de formation, de telle sorte qu'il soit un support et un levain favorable à une mise en commun cultivant les ferments du collaboratif et les potentiels de chacun. On peut considérer que cette zone complexe d'élaboration du sens, médiatisée par l'évaluation, réinvente une forme individuelle et collective de maïeutique et d'empowerment, permettant tout à la fois à chacun de repartir de soi-même enrichi par la confrontation à des altérités et de co-construire de nouveaux possibles.

Les dynamiques d'intersubjectivité renvoient à des positions différenciées et à la distance entre des partenaires. L'enjeu du pouvoir est aussi crucial : l'EEE tendant à donner un pouvoir accru à l'étudiant, elle peut contribuer à réduire le pouvoir de l'enseignant tant sur la dimension symbolique que sur le plan institutionnel. Elle peut ainsi être perçue comme une menace pour l'autorité : il y a aussi un fort enjeu autour du pouvoir et de l'autorité dans l'évaluation.

4.2 - Recherche collaborative sur les pratiques d'évaluation soutien d'apprentissage : l'évaluation subjectivante

L'évaluation subjectivante ne renvoie pas seulement à des états d'âme subjectifs, mais à la façon dont les différentes parties prenantes de l'évaluation sont engagées dans leurs manières de comprendre et de se situer affectivement, cognitivement et activement dans l'évaluation. Dans le cadre de l'universitarisation de la formation au métier d'enseignant dans les premier et second degrés, désormais traduite en master, la mise en place de travaux de recherche co-encadrés par des enseignants-chercheurs et des formateurs de terrain a fourni le terreau institutionnel propice à la construction d'une démarche collaborative déployée successivement entre 2010 et 2013 dans une recherche sur les pratiques évaluatives à l'école primaire, puis de 2017 à aujourd'hui dans le secondaire.

Le risque aliénant des pratiques d'évaluation traditionnelles, basées sur le classement par la note, a été largement documenté par des décennies de recherches en sciences sociales, aussi bien en termes de reproduction et de légitimation des inégalités qu'en termes identitaires, motivationnels et d'apprentissage (Butera, Buchs et Darnon, 2011). L'évaluation aliénante ou « nocive » est celle qui, en catégorisant l'élève, tend à le positionner parmi d'autres, dans une logique de comparaison sociale, au détriment d'un projet de progression propre, caractéristique qui avait déjà été mentionnée pour les enseignants dans le cas de l'EEE. A contrario, une évaluation qui a une visée émancipatrice tend à envisager l'élève comme un sujet-apprenant en tant que tel, capable de se situer et de participer à une stratégie de progression personnelle. Il s'agit, par le moyen de l'évaluation, de lui faire prendre conscience que son histoire s'inscrit dans une dynamique singulière dans laquelle ses actions sont productrices d'effets sur soi et sur autrui. C'est ce que nous nommons une évaluation subjectivante qui ne se réduit pas à l'autoévaluation, ou plus généralement à l'implication de l'élève dans l'évaluation mais renvoie à la création d'un ordre intersubjectif entre l'enseignant et l'élève. Dans cet ordre intersubjectif, la construction dialoguée des jugements est centrale et requiert l'encouragement, l'étayage et la valorisation de la parole de l'élève.

Nous avons conduit une recherche collaborative sur l'explicitation par des PEMF (professeurs des écoles maitres formateurs) de leurs pratiques évaluatives des apprentissages des élèves (Younès, Sasse et Darj, 2016). Ces enseignants, eux-mêmes formateurs d'enseignants, ont pour caractéristique de déclarer promouvoir des pratiques que nous pouvons associer à une « évaluation soutien d'apprentissage » (Allal et Laveault, 2009). La recherche, initialement descriptive, s'est progressivement recentrée autour de l'objectif de co-construire avec eux ce qui pourrait relever d'une évaluation des élèves émancipatrice et subjectivante dans leurs pratiques quotidiennes d'évaluation, sans toutefois leur proposer une modélisation préalable de cette notion qui s'est construite au cours de l'investigation. A savoir une évaluation à même de mobiliser les propres capacités de jugement d'un individu qui est le produit d'une histoire dont il est le sujet en devenir.

Nous présentons ci-dessous les différents cadres théoriques que nous avons convoqués, ainsi que la démarche de recherche et les résultats co-construits. Ce faisant, nous insistons sur la question du sens partagé dans le collectif collaboratif quant à la part de l'élève dans le processus d'évaluation soutien d'apprentissage (ESA), afin d'explorer ce que nous avons nommé des « strates de la subjectivation ». Initialement centrée sur la dimension temporelle de la séquence pédagogique, cette étude s'est progressivement focalisée sur les dispositifs de subjectivation /

intersubjectivation à l'œuvre dans les pratiques d'évaluation comme dans la recherche collaborative.

4.2.1. Cadre théorique de l'enquête : subjectivation / intersubjectivation et évaluation soutien d'apprentissage

Fondée sur une alliance entre chercheurs et praticiens pour construire ensemble les connaissances liées à leur pratique de l'ESA, l'enquête (au sens de Dewey, 1993) dont il est rendu compte ici s'inscrit dans le courant qualitatif des recherches collaboratives sur l'enseignement (Cole, 1989; Cole et Knowles, 1993; Desgagné, 1997, 2007; Mottier Lopez et al. 2010; Morrissette, Mottier Lopez et Tessaro, 2012). Dans cette approche, un primat est accordé au sens que le praticien construit à propos des situations dans lesquelles il agit. Et ce en s'appuyant sur son jugement reconnu comme l'expression d'une « compétence d'acteur en contexte » au sens de Giddens (1987), à savoir: « ... sa capacité à exercer son jugement et à orienter sa prise de décision en situation, en fonction des conditions qui prévalent et des enjeux en présence, dans le contexte d'action spécifique qui est le sien » (Desgagné, 1997, p. 373). Cette approche s'appuie sur deux hypothèses théoriques principales mettant en jeu subjectivation et intersubjectivation qui ont orienté les choix méthodologiques de l'analyse empirique :

- La pratique de chaque enseignant est idiosyncrasique. Il s'agit de l'expression d'un mode de connaissance personnelle et professionnelle formée et informée par les événements et les expériences à la fois passés et présents situés dans leurs différents contextes, personnels, institutionnels, culturels et sociaux (Cole and Knowles, 1993). S'avère déterminante, la part de cette subjectivation trans-temporelle; processus par lequel le sujet se construit en fonction de strates d'expériences acquises au fil du temps, d'une adaptation aux milieux et d'un projet en devenir.
- La mise en relation constante des différents points de vue des pédagogues et des chercheurs est productrice sur un double plan. D'une part, elle crée un espace de reconnaissance et de coopération propre à favoriser une forme de nouvelle subjectivation grâce à ce processus d'intersubjectivité où le sujet s'envisage à l'autre. D'autre part, elle est productrice d'interfécondation entre le champ de la pratique et le champ scientifique, conduisant à la fois au développement professionnel des praticiens et à celui des connaissances scientifiques (Desgagné et Berdnarz, 2005). Cette intersubjectivité professionnelle permet ainsi une co-construction du sens des pratiques.

- ***Évaluation formative et évaluation soutien d'apprentissage***

En lien avec les théorisations des précurseurs tels Allal, Bonniol, Cardinet, Hadji ... la « perspective élargie » de l'évaluation formative définie dans Allal et Mottier Lopez (2005) se veut au service de

l'apprentissage des élèves. Elle met l'accent sur une double fonction de régulation, celle de l'activité des élèves et celle de l'action pédagogique des enseignants, notamment dans le sens d'un ajustement différencié aux caractéristiques et aux besoins des élèves. Dans le même ordre d'idée, l'ESA est considérée comme « faisant partie des pratiques quotidiennes des élèves et des enseignants qui, individuellement et en interaction, recherchent, réfléchissent sur et réagissent à l'information provenant d'échanges, démonstrations et observations afin de favoriser les apprentissages en cours (Allal et Laveault, 2009, p.102). Dans cette définition, il est rendu compte de pratiques visant un meilleur équilibre des rôles entre enseignant et élèves et le renforcement des échanges entre eux. L'ESA est la traduction proposée par Allal et Laveault (2009) des concepts anglosaxons *Assessment as et for Learning*. D'après ces auteurs, ce concept, apparenté à celui d'évaluation formative, le dynamise en soulignant l'engagement actif de l'élève dans ses apprentissages et l'élargit en estompant dans certains cas, la distinction classique entre évaluation formative et sommative, considérant que cette dernière peut également, dans certaines conditions, contribuer à soutenir l'apprentissage (Allal et Laveault, 2009, p. 99). Dès le départ, les travaux sur l'évaluation formatrice avaient mis l'accent sur le rôle nécessairement actif de l'élève dans le processus d'évaluation, en particulier dans l'appropriation des critères, dans la maîtrise des processus d'anticipation et de planification des actions, dans l'autocontrôle ainsi que dans l'autocorrection des erreurs (Nunziati, 1990). Plus récemment l'ESA rend compte de pratiques visant à un partage plus équilibré de la responsabilité de l'élève et de l'enseignant dans le processus d'évaluation. Selon cette conception, les rôles de chacun de même que les processus d'évaluation et d'apprentissage apparaissent davantage liés et moins séparés (Laveault, 2012). Laveault (2007) va jusqu'à proposer de concevoir l'évaluation comme un système plus ou moins régulé par l'élève selon le degré de contrôle exercé sur la définition des buts et des standards, sur le feedback et sur la remédiation. Il représente l'évaluation sur un continuum variant d'une régulation assumée par l'enseignant à une régulation assumée par l'élève. Pour Earl (2003), dans la mesure où l'enjeu consiste à ce que l'élève soit un acteur critique de son apprentissage, l'autoévaluation devient une compétence à développer. Les régulations à l'œuvre dans le processus d'ESA étant cognitives mais également sociales et émotionnelles (Cosnefroy, 2011), c'est l'ensemble du processus qui est à considérer comme étant socio-affectivo-cognitif. Il s'agit là d'activer la rencontre entre jugements subjectif et intersubjectifs.

Dans une perspective d'évaluation formative élargie (Allal et Mottier Lopez, 2005), située et contextualisée (Mottier Lopez, 2008, 2012) cette recherche visait à participer à la construction d'un savoir intersubjectif sur les pratiques d'ESA en classe, produit dans une logique de coopération entre chercheur et enseignants.

4.2.2. Les trois étapes de la recherche collaborative

Trois grandes étapes ont structuré cette recherche collaborative : identifier les pratiques d'ESA, repérer les dynamiques de subjectivation et mener un retour réflexif.

Etape 1 - identifier les pratiques d'évaluation soutien d'apprentissage

Cette recherche collaborative à visée compréhensive et critique a été menée dans l'académie de Clermont-Ferrand durant 3 années (2010-2011; 2011-2012; 2012-2013) avec 4 maitres-formateurs déclarant inscrire leurs pratiques de classe dans une orientation d'évaluation soutien d'apprentissage. Nous travaillons ici sur leurs pratiques d'évaluation des élèves dans leurs classes. Leur participation a été requise tout au long de ces trois années dans l'élaboration des connaissances produites sur leurs pratiques en cherchant à décrire au mieux ce qu'ils font à partir d'une démarche inductive. Cette posture a orienté le déroulement de l'investigation en nécessitant que les enseignants s'engagent non seulement à expliciter leurs manières de faire, mais également à les mettre en débat et à les formaliser. De son côté, le chercheur a accepté que le projet théorique initial se redéfinisse au fil de l'élaboration collective.

Au départ, le projet était d'identifier les pratiques d'ESA de ces quatre enseignants expérimentés. La réflexion visait, d'une certaine manière, à s'appropriier cette conception, au regard du cadre théorique de l'évaluation formative, bien connu dans le monde francophone. Ce cadre théorique n'a cependant été que peu évoqué avec les enseignants au démarrage de la recherche, pour deux raisons. D'une part, il s'agissait de saisir et de restituer l'explicitation des pratiques des enseignants collaborant au plus près de leur « réalité » perçue, en évitant au maximum d'induire des phénomènes de désirabilité sociale. D'autre part, nous souhaitions instaurer une relation égalitaire et ne pas accorder de primat à la théorie sur la pratique mais au contraire mettre l'accent sur un enrichissement mutuel de l'une et de l'autre. Le cadrage initial retenu s'est limité à situer l'évaluation dans une chronologie. Il s'agissait en effet d'identifier l'ESA dans sa temporalité pédagogique et d'en énoncer les différents moments: en amont d'une séquence, pendant celle-ci et en aval. Le flou autour du sens même de ce cadre en tant que tel n'a pas été éclairci par le chercheur volontairement afin d'amorcer une démarche d'élaboration théorique non prédéfinie par les concepts de la littérature spécialisée.

Le premier temps a été consacré à une activité réflexive et partagée sur les pratiques des uns et des autres à partir de ce cadre commun (en amont, au cours, en aval de la séquence). Il a tout d'abord été procédé à des entretiens croisés suivis de rédactions individuelles. Dans la séance de travail suivante, les productions individuelles ont été remises en discussion sous forme d'échanges, de commentaires et de débats – consignés dans un journal de recherche – avant que les écrits soient repris et communiqués au groupe. Ces descriptions ont alors été mises en regard

les unes par rapport aux autres avec l'objectif de faire émerger les points d'accord, les divergences, les questionnements, également consignés. Suite à cela, une schématisation de chacune des quatre pratiques a été réalisée individuellement par son auteur à partir du cadre temporel identifiant la préparation en amont, l'évaluation au cours de la séquence d'apprentissage et le suivi en aval de l'évaluation. Au fil des discussions, des éléments théoriques, introduits par l'un ou l'autre des participants, ont été explicités. On en trouve la trace dans les discours et leur structuration tels qu'ils sont donnés à voir dans les analyses suivantes.

- **Analyse des procédures d'ESA mises en œuvre**

Il apparaît que ces schématisations procédurales présentent des similitudes et des spécificités. Les éléments structurants partagés sont de divers ordres: explicitation du cadre de l'évaluation, des objectifs d'apprentissage et des critères, instauration d'un climat motivationnel plutôt orienté vers la maîtrise, articulation et cohérence de l'évaluation et de l'apprentissage, différenciation selon différentes formes, inscription dans un parcours de progression de l'élève. Chez ces quatre enseignants, il y a adaptation et personnalisation d'outils et de dispositifs visant à répondre in situ à ces différents paramètres.

Eléments structurants partagés entre les 4 enseignants

<p><i>En amont de la séquence</i></p> <ul style="list-style-type: none"> - Information des parents, élèves, collègues du cadre de l'évaluation: sens, dispositif, valeurs - Instauration d'un climat motivationnel d'apprentissage sécurisant à maintenir au fil du temps - Définition des objectifs d'apprentissage et de maîtrise - Définition, explicitation des critères de réalisation et des critères de réussite - Choix de tâches adaptées aux niveaux des élèves
<p><i>Pendant la séquence</i></p> <ul style="list-style-type: none"> - Observation des élèves, individuellement et en interaction - Recherche du sens des informations: échanges, comparaison (des réalisations), argumentation... - Auto-évaluation - Feedback en situation - Ajustements
<p><i>En aval de la séquence</i></p> <ul style="list-style-type: none"> - Bilan en termes de réussites et de difficultés - Plan d'action: adaptation à la zone proximale de développement - Appréciation de la progression en direction des enseignants, des élèves et des parents

- **Tensions perçues entre contextes et valeurs personnelles**

Les débats autour de ces schématisations des procédures d'évaluation mettent en évidence les tensions issues des divergences ressenties entre les différentes cultures pédagogiques, sociétales,

institutionnelles et personnelles, au croisement desquelles les enseignants exercent leur métier. Quatre thèmes principaux qui illustrent ces tensions ont été dégagés.

- *Entre cultures du milieu pédagogique et de la société*

Ces enseignants accordent une importance de premier ordre aux parcours d'apprentissage de leurs élèves et donc à leurs progrès. L'évaluation est au service de la conscientisation de l'élève et de la régulation pédagogique (différenciation des aides). Ils considèrent cette culture pédagogique comme opposée à la culture de la compétition et des résultats, caractéristique de la société actuelle et dont sont imprégnés les parents, en particulier des élèves en réussite, ainsi que ces élèves eux-mêmes. Deux enseignants sur les quatre ont conservé la notation dans leurs pratiques d'évaluation pour satisfaire cette demande sociale (une troisième enseignante a également conservé la notation considérant qu'elle offre plus de nuances et surtout qu'elle prépare mieux l'élève à l'avenir des notes qu'il attend). Ils sont également dubitatifs quant à l'évaluation des compétences telle qu'elle est prescrite dans les livrets et les évaluations nationales. Ils considèrent d'une part que ces livrets sont peu lisibles et peu utiles. Selon une enseignante, ils pourraient même avoir un effet délétère « d'étiquetage » des élèves et d'autre part l'évaluation binaire des compétences en acquis/non acquis ne reflète pas la réalité du processus d'apprentissage plus progressif que ne l'indique une telle dichotomie.

- *Entre cultures de l'école et du collège*

Ces quatre enseignants pratiquent des formes de différenciation pédagogique c'est-à-dire qu'ils adaptent leur enseignement aux besoins et caractéristiques individuels des élèves. Cette culture professionnelle est perçue comme opposée à la culture uniformisante du collège où les élèves seront soumis à un enseignement semblable indépendamment de leurs besoins variés. Ils considèrent que cette culture du collège tout à la fois anonymisante et classante constituera un milieu délétère pour les élèves les moins avancés qui seront alors en échec.

- *Contextes particuliers d'exercice professionnel*

L'impact des contextes d'établissements dans lesquels ces enseignants exercent est apparu au niveau de l'utilisation des notes et de la différenciation pédagogique. Une seule personne parmi les quatre enseignants a totalement abandonné les notes au profit de l'utilisation d'une échelle d'évaluation des compétences en 5 points. Il s'agit d'une décision collective, partagée par tous au sein de l'école. En ce qui concerne la différenciation pédagogique, si tous ces enseignants la pratiquent d'une manière ou d'une autre, elle est beaucoup plus généralisée dans un cas, jusqu'à l'inclure dans l'évaluation

sommative (avec des épreuves différentes) puisque certains élèves ont des exercices correspondant à des niveaux CE1 (2ème année primaire) alors qu'il s'agit d'une classe de CM1 (4ème année primaire). Cette façon de faire a été le résultat d'un ajustement à un milieu particulier : l'école accueillant beaucoup d'enfants du voyage, les différences sont rendues saillantes à tel point qu'un enseignement non différencié est considéré impossible par l'enseignante.

- *Divergences inter-individuelles*

Enfin, bien que ces enseignants partagent de nombreuses valeurs, des divergences de positionnement vis-à-vis de l'évaluation sont également repérables en ce qui concerne l'emploi des notes, l'intérêt de la compétition et les bilans de fin de séquence. Les débats les plus vifs ont porté sur la notation, chiffrée ou par compétence, reflétant la polémique actuelle traversant le monde de l'éducation, les partisans de la note chiffrée en soulignant l'aspect précis, lisible et motivant pour certains élèves alors que ses détracteurs relèvent au contraire son caractère démotivant et favorisant une comparaison sociale dévalorisante pour une partie des élèves. Néanmoins les enseignants utilisant la note chiffrée pratiquent tous un système mixte (note et repérage d'un niveau d'acquisition de la compétence).

Étape 2 : centration sur les dynamiques de subjectivation dans les pratiques d'ESA

À l'issue de cette première année consacrée à travailler sur l'ESA dans son épaisseur chronologique et à mettre en évidence les valeurs collectives et individuelles qui s'y rattachent, un pas théorique a été réalisé en décidant de se centrer sur les dynamiques de subjectivation, vues comme des dimensions transversales dans les pratiques d'ESA. En effet, au fil des échanges, le groupe a considéré important de pouvoir sortir d'un modèle de dépendance vis-à-vis d'une évaluation essentiellement portée par autrui (en l'occurrence celle de l'enseignant). Ainsi, l'idée de subjectivation a émergé, afin d'inscrire les pratiques évaluatives dans une dynamique d'appropriation et de transformation centrée sur un parcours singulier et sur la reconnaissance d'un sujet auteur (Jorro, 2007).

Au cours de cette étape, le chercheur a construit et soumis un cadre théorique de ce que pourrait être une « évaluation subjectivante ». Très vite, le collectif de recherche a estimé erroné de croire que l'ensemble du processus serait uniquement sous le contrôle de l'enseignant et celui de l'élève alors qu'une grande part des dynamiques liées à l'apprentissage leur échappe à l'un comme à l'autre, étant donné le poids des contextes, des histoires singulières et des processus de subjectivation, puisque le sujet se construit à travers cette histoire personnelle. C'est pourquoi il a

été considéré que cette orientation devait être envisagée comme une ouverture de possibles et non comme un modèle normatif. Le collectif collaboratif s'est alors donné comme objectif de repérer les situations subjectivantes dans les schématisations d'ESA, étant donné l'importance qu'il a choisi de leur accorder.

- **Repérage de situations subjectivantes**

Les situations évaluatives considérées comme les plus susceptibles de participer au processus de subjectivation de l'élève ont été repérées aux trois principaux niveaux du processus d'ESA par chacun des enseignants. Les indices de repérage ont été définis comme étant les situations qui permettent le plus à l'élève d'être en position d'auteur dans le processus d'évaluation.

Tableau de repérage des situations subjectivantes

En amont

- Les quatre enseignants accordent un soin particulier à la *préparation des bilans* de fin de séquence. Explicitement référés aux compétences attendues, ces bilans portent spécifiquement et uniquement sur celles-ci et sont positionnés après un entraînement approfondi et répété. Les enseignants veillent à ce que les élèves soient des acteurs de cette préparation en l'ajustant à leurs capacités d'anticipation. Au niveau du CP, il s'agit de pré-bilans à travers lesquels les élèves sont amenés à faire le point avec l'enseignant sur leurs acquis et leurs manques. Au niveau du CM2, les élèves sont renvoyés à leurs productions antérieures.
- Le *choix par l'élève du niveau de l'exercice de renforcement* qui lui convient le mieux est pratiqué dès le CP par l'un des quatre enseignants. Bien que l'enseignant se réserve la possibilité, après concertation, de demander à l'élève d'essayer de choisir un niveau qui lui correspondrait mieux, les élèves restent maîtres de leur choix et peuvent le modifier s'ils le jugent finalement trop difficile (ou trop facile).

Au cours de l'évaluation

- La pratique de l'*autoévaluation* est généralisée avec des modalités différentes selon les classes et selon les activités, notamment au niveau des « arrangements des rôles » enseignant/élève/élèves (Laveault, 2007). L'analyse des stratégies d'exécution des tâches ou de résolution de problème ainsi que des erreurs est également généralisée.

Dans le suivi en aval

- Le *suivi longitudinal des compétences* génériques et des sous-compétences¹¹ est un axe fort de la pratique de 2 enseignants. Pour l'un de ces enseignants, il s'agit d'une présentation avec un code couleur facilitant le repérage par les parents, les élèves et l'enseignant, de l'évolution aussi bien des compétences génériques que des sous-compétences. Pour une autre enseignante, les élèves consignent eux-mêmes leurs progrès sur leur « chemin d'apprentissage ».

Chacune des étapes repérées dans le tableau ci-dessus contribue à l'étayage des strates de subjectivation déterminées à la fois par des enchainements procéduraux et par un climat d'ensemble. Lors de l'analyse par le collectif collaboratif des situations évaluatives subjectivantes repérées par chacun des enseignants, il est apparu qu'une forme de réduction de l'idée d'évaluation subjectivante avait été effectuée du fait de son assimilation à la seule implication manifeste de l'élève. Cette réduction a conduit à omettre des situations dirigées par l'enseignant alors propres à éclairer le sens de la pratique d'ESA pour les élèves et pour leurs parents.

Étape 3 : retour réflexif sur la recherche collaborative

Des entretiens semi-directifs croisés ont été menés portant sur le bilan et les apports de la recherche expérimentée, sur l'origine et la genèse des formes actuelles des pratiques et sur les évolutions envisagées. Ces entretiens, enregistrés, ont été suivis par la rédaction individuelle, à froid, d'une synthèse. Une analyse de contenu des entretiens et des écrits individuels a ensuite été effectuée.

La part de l'histoire personnelle dans la genèse des pratiques actuelles apparaît déterminante dans chaque trajectoire professionnelle des enseignants. Cet ancrage biographique que nous n'avons fait qu'effleurer constitue une piste de recherche à poursuivre étant donnée son importance. Dans les 4 cas, il est fait référence à un vécu d'enfant, d'élève, d'étudiant avec un fort accent mis sur la transmission de valeurs, que ce soit par une mère institutrice « qui faisait du pseudo-Freinet », par un père enseignant dans le technique qui « transmettait les valeurs de l'effort à ses élèves » ou en tant que « fille d'immigrés qui avait compris l'importance de la maîtrise de la langue ». Si les collègues semblent avoir eu peu d'influence sur la construction des gestes professionnels, la formation initiale et continue est largement évoquée à ce niveau. L'influence de la formation initiale est mentionnée dans le cas d'une formation liée à la pratique de l'enseignement alors que dans le cas d'une formation initiale d'ingénieur, c'est la formation

¹¹ La compétence par exemple la maîtrise de la multiplication se décline en sous-compétences comme connaître les tables de multiplication, savoir effectuer un calcul posé, maîtriser la distributivité de la multiplication sur l'addition.

continue qui est remémorée. Quels que soient leurs parcours, ces quatre enseignants ont mis l'accent sur une évolution, des ajustements permanents et une quête adaptative d'outils satisfaisants sur le plan de leurs finalités comme sur le plan pragmatique.

Du point de vue des maitres formateurs, **les apports de la participation à la recherche** sont déclarés notables. Ils se déclinent selon quatre thèmes.

- *Une meilleure compréhension du sens de sa pratique*

Un nombre important de verbatims témoigne d'une meilleure compréhension du sens de sa pratique en liaison avec l'articulation menée entre pratique/théorie/pratique (Altet, 1996), réalisée grâce au travail de réflexivité, d'intersubjectivité et de conceptualisation. Les occasions de présentation, structuration, argumentation, discussion des pratiques de chacun et les mises en relation avec des connaissances issues de la recherche ont été jugées fructueuses. Le positionnement professionnel spécifique de ces enseignants expérimentés reconnus pour leurs compétences professionnelles, et investis dans la formation des jeunes professeurs des écoles, les a amené à attendre et à valoriser les savoirs scientifiques, qu'ils arrivent à référer à leur univers professionnel et qui participent à en ancrer le sens. La recherche collaborative a permis non seulement une appropriation de l'idée d'évaluation subjectivante, mais a contribué également à nourrir la conceptualisation émergente de celle-ci.

- *Une légitimation de sa pratique*

Le processus de réflexion, de formalisation et de conceptualisation a conduit également à une forme de légitimation de la pratique d'évaluation, liant activité métacognitive et motivation, climat de classe et coopération.

- *Un développement professionnel en tant que maître formateur*

Un autre apport de cette recherche consiste en un développement de sa capacité à agir dans son champ professionnel en tant que maître formateur. Le rôle des maitres formateurs dans la formation et l'accompagnement des jeunes enseignants ainsi que dans leurs échanges avec leurs pairs apparait ainsi renforcé et irrigué par l'articulation pratique/théorie construite au cours de cette recherche. La légitimation des pratiques d'évaluation produite par le partage et la théorisation associée à une capacité accrue à les expliciter et à les communiquer est considérée comme un axe de développement professionnel important.

- *Un développement professionnel en tant que professeur des écoles*

Des évolutions dans les pratiques d'évaluation en classe sont également mentionnées. Ces évolutions consistent principalement à accroître les échanges entre élèves et entre

élève et enseignant, une forme d'écho à la conceptualisation en chemin de l'évaluation subjectivante et de la méthodologie de la recherche collaborative.

4.2.3. Un pas vers une évaluation subjectivante et non seulement objectivante

Les procédures d'évaluation déterminées dans les quatre cas analysés s'avèrent être des axes forts de la pratique, qui s'instancient de manière située et qualitative de façons diverses. Chaque enseignant a développé une pratique personnelle à travers des dispositifs finalisés ajustés en fonction du niveau de la classe, du contexte de et dans l'école, et de son histoire personnelle. Au-delà des différences, ils partagent la visée de rendre l'élève plus actif dans sa dynamique d'apprentissage en s'appuyant sur l'évaluation soutien d'apprentissage. Demander à l'élève de choisir le niveau d'exercice qui lui convient, d'exprimer un avis critique sur les tâches qu'il effectue et sa façon de le faire, et de participer au suivi de ses progrès, c'est l'impliquer autrement en lui conférant un pouvoir d'action conséquent. A l'opposé d'une évaluation qui objective l'élève en le classant, en le catégorisant et en ne le faisant pas progresser, il est considéré que ces enseignants tendent à s'inscrire dans une autre orientation, celle d'une « évaluation subjectivante » qui considère l'élève comme un sujet capable de comprendre l'environnement dans lequel il se trouve, de se situer dans ses apprentissages et de s'impliquer dans une dynamique de progression. Ce type d'évaluation, par l'importance accordée aux dynamiques de subjectivation et d'intersubjectivation s'éloigne de toute tentation positiviste concevant l'évaluation comme un système seulement objectivant. Les « strates de la subjectivation » qui s'élaborent autour de la création d'une dynamique d'évaluation entre le maître et les élèves, plus collaborative quant au sens, cadre, objectifs et critères de l'évaluation et par rapport à son inscription dans une dynamique de progression personnelle, enrichissent les modalités de régulation au cœur de la littérature sur l'ESA. Cependant, la situation de la classe dans son ensemble, et son climat, participent fortement à cette élaboration.

Cette recherche collaborative manifeste la fécondité d'une réflexion en mouvement qui s'est transformée et métamorphosée dans le cours du travail et de la coproduction de connaissances. Elle a constitué un atelier d'intersubjectivité durant lequel les participants ont pris en considération la pensée de l'autre dans leur propre jugement et ont été capables de reconnaissance réciproque dans un espace de dialogue critique. C'est cet espace d'intersubjectivité qui a créé les conditions d'affirmation de la posture qu'Habermas appelle la subjectivité critique.

4.3 - Co-construction d'un dispositif d'évaluation de l'enseignement à visée synergique

L'étude, en contexte, des dynamiques évaluatives subjectives et intersubjectives a été poursuivie dans une autre recherche collaborative portant sur l'expérimentation d'un dispositif coopératif d'évaluation de l'enseignement, associant enseignants et chercheur-médiateur dans la construction et l'analyse des données de l'évaluation (Younès et Paivandi, 2017). Ont été étudiés, les effets d'une réflexion partagée entre enseignants et chercheur-médiateur, à propos de l'ajustement d'un dispositif d'EEE intégrant une confrontation avec le point de vue des étudiants. Nous étudions en quoi la micro situation créée à partir de l'invention d'un dispositif situé dans le cadre d'une formation de master préparant au métier d'enseignant permet d'explicitier les dynamiques intersubjectives et coopératives en jeu entre enseignants, étudiants et chercheur-médiateur, dans le processus et les effets de l'EEE.

4.3.1. Une orientation subjectivante/intersubjectivante située du dispositif en jeu

Nous avons souligné l'importance d'une « écologie des pratiques » de l'EEE, à savoir prenant en compte les caractéristiques et dynamiques propres à chaque milieu, considéré en tant qu'environnement d'étude et de cultures universitaires aussi bien disciplinaires que pédagogiques. La prégnance des écosystèmes locaux apparaît nettement à travers le littérature (Romainville et Coggi, 2009 ; Younès et Romainville, 2012). Les objectifs de la formation, la culture disciplinaire ainsi que les conditions d'enseignement définissent un milieu spécifique orientant le sens du processus d'évaluation.

Le dispositif d'évaluation expérimenté a été construit en tant que médiation psycho-socio-technique en référence avec la théorie des dispositifs. Il a « une fonction de support, de balise, de cadre organisateur à l'action. Il procède essentiellement à des mises en ordre qui soutiennent l'action de l'individu, il crée des effets de signification qui procurent des ressources pour un auto-pilotage. Mais si le dispositif organise et rend possible quelque chose, il n'en garantit cependant pas l'actualisation. Il fait simplement exister un espace particulier préalable dans lequel ce « quelque chose » peut se produire, ce qui selon certains auteurs actualise les potentialités d'un milieu en favorisant les relations entre les différents éléments du milieu et la construction de significations » (Peeters et Charlier, 1999, p.19).

La recherche dont il est rendu compte s'attache à décrire ce dispositif qui s'est voulu situé et collaboratif dans une visée synergique, formative et transformative. Il est patent que, dans de nombreux cas, les dispositifs consistant à administrer aux étudiants, généralement une fois par année ou par semestre, un questionnaire standardisé dont les résultats sont communiqués notamment aux enseignants, ne produisent pas systématiquement les effets d'amélioration

escomptés (Marsh, 2007 ; Edström, 2009). Ces dispositifs sont d'ailleurs souvent peu investis par les étudiants et les enseignants comme en témoignent les faibles taux de participation des premiers et d'engagement des seconds (voir à ce sujet les nombreux comptes-rendus d'expériences institutionnelles et de recherche que l'on peut trouver sur les sites des universités ou dans des publications (ex, Heyde et Le Douris, 2009 ; Quinton, 2009 ; Beran et Rokosh, 2009 ; Burden, 2008). Ils peuvent se trouver réduits à une procédure bureaucratique vidée de son sens tout en contribuant à façonner la culture de l'évaluation d'une organisation (Dahler-Larsen, 2005). Au contraire dans le dispositif coopératif coproduit, les mises en relation des différents points de vue des étudiants et des enseignants qui passe par l'expérience du désaccord entre les acteurs - entre les étudiants, entre l'enseignant et les étudiants et entre enseignants- est centrale. Il a été considéré plus particulièrement en quoi ces mises en relation peuvent être productrices de sens dans la mesure où elles créent un espace dialogique de communication et de coopération.

4.3.2. Présentation du dispositif d'évaluation de l'enseignement dans un master de formation au métier d'enseignant

Le dispositif d'évaluation a été expérimenté dans le contexte d'une formation de master au métier d'enseignant dans un Institut Universitaire de Formation des Maitres (IUFM). L'IUFM (aujourd'hui ESPE, école supérieure du professorat et de la formation) assure essentiellement des formations de niveau master dans le domaine de l'enseignement et de la pédagogie. En France, le master préparant au métier d'enseignant du premier degré, qui se déroule en deux années, a la particularité d'être articulé à une politique nationale en matière de recrutement d'enseignants dans les premiers et seconds degrés. Ainsi entre le master 1 et le master 2, l'étudiant passe un concours national. La formation en alternance en deuxième année de master est rémunérée en cas de réussite au concours. Les impératifs du concours en master 1 et de l'entrée dans le métier en master 2 sont donc les référents fondamentaux du parcours de formation, des référents pouvant entrer en contradiction dans le curriculum. Il est à noter que la politique d'évaluation de l'enseignement de l'université est encore en construction et en émergence dans cet IUFM, même si, comme partout, certains enseignants en ont une pratique personnelle et qu'elle se développe dans le cadre des comités de perfectionnement des formations et d'un pilotage central.

Les acteurs de l'expérimentation sont : une enseignante-chercheuse en sciences de l'éducation qui a le rôle de chercheur-médiateur et 4 enseignants motivés pour tenter l'expérience. L'enseignant-chercheur est relayé par deux étudiants engagés dans la recherche via la réalisation de leur mémoire de master. L'enseignant-chercheur médiateur est un collègue des enseignants participants, identifié comme poursuivant ses recherches sur la thématique, membre de la commission évaluation de l'université et sollicité par son établissement pour contribuer à / et

soutenir la réflexion sur le sujet mais sans mission officielle. Les enseignants sont : un enseignant de physique intervenant avec les étudiants de première année de master, une enseignante de français intervenant également avec les master 1, deux enseignants de français intervenant avec les étudiants de master en deuxième année. Ces travaux se sont déroulés au cours de l'année 2012/2013.

- ***Cinq étapes dans ce dispositif d'évaluation de l'enseignement***

- *Etape 1 : Coconstruction entre enseignant et chercheur/médiateur*

De premières rencontres ont eu lieu entre l'enseignant-chercheur-médiateur et chacun des enseignants séparément. Cette première rencontre a discuté les objectifs, le cadre de la démarche mise en œuvre, l'enseignement dans le contexte de la formation, et la co-construction d'un dispositif d'EEE adapté au projet pédagogique de l'enseignant. L'instauration d'un climat de confiance était déterminante. Le chercheur-médiateur a proposé une trame de questionnaire issue des dimensions validées du SEEQ de Marsh (1982) à savoir les apprentissages, l'enthousiasme de l'enseignant, la structuration et l'étendue de l'enseignement, les attentes vis-à-vis des étudiants, la clarté des attentes et des modalités d'évaluation des étudiants, la qualité des interactions suscitées entre les étudiants, la relation entre l'enseignant et l'étudiant. Cette trame a été ajustée et spécifiée à chaque enseignant en fonction des objectifs et des modalités particulières de l'enseignement en question. Il a également été décidé des moments opportuns pour le recueil des points de vue des étudiants et de leurs modalités.

- *Etape 2 : Recueil quantitatif et qualitatif des points de vue des étudiants*

Les points de vue des étudiants sur les enseignements ont été recueillis en l'absence de l'enseignant sur un moment de cours soit par le chercheur-médiateur soit par les étudiants engagés dans la recherche. Le recueil quantitatif à l'aide de boîtiers de vote a été complété par les commentaires écrits des étudiants sur les formulaires distribués sous forme de documents papier.

- *Etape 3 : Co interprétation des EEE*

Une synthèse des résultats quantitatifs et qualitatifs a été effectuée par le chercheur médiateur ou les étudiants puis soumise à l'analyse de l'enseignant lors d'un second entretien, au cours duquel l'enseignant a commenté les résultats des EEE.

- *Etape 4 : Mise en débat entre enseignants et étudiants*

Suite à cette analyse, l'enseignant a organisé un temps de restitution commentée et discutée des résultats de l'EEE avec les étudiants en présence du chercheur médiateur ou des étudiants-chercheurs.

- *Etape 5 : Bilan et mise en projet*

Enfin un troisième entretien a été consacré à faire un retour à froid sur la démarche d'évaluation dans sa globalité. Il s'agissait d'étudier comment les enseignants l'avaient vécue, quel regard ils portaient sur les points de vue des étudiants, comment ils les comprenaient, et quelles suites ils envisageaient dans leurs enseignements. Cet entretien s'est déroulé selon deux modalités : individuelle dans le cas du professeur de physique, collective pour les trois enseignants de français.

4.3.3. Trois études de cas

Trois études de cas ont été conduites au cours de l'évaluation et après pour tenter d'appréhender les phénomènes liés au processus d'évaluation, en leur contexte, celui d'une formation située, professionnalisante dès l'origine, universitarisée depuis peu, et fortement contrainte par sa double visée de préparation au concours et au métier. Ce mode d'enquête à visée d'analyse et de compréhension consiste à étudier en détail et avec des sources d'information diverses, un phénomène dans son contexte, tel qu'il s'est déroulé dans une situation particulière, réelle ou reconstituée, jugée représentative de l'objet à étudier (Yin, 1984). Elle s'inscrit dans le courant des recherches qualitatives dont l'orientation interprétative et compréhensive donne une large place à la subjectivité, afin de comprendre les significations que les individus donnent à leur propre vie et à leurs expériences ; des significations et interprétations élaborées par et dans les interactions sociales (Anadon, 2006).

La méthodologie des études de cas s'est appuyée sur plusieurs outils mobilisés à chaque étape du dispositif : trois entretiens enregistrés et retranscrits conduits en amont et en aval du recueil des points de vue des étudiants, les données quantitatives et qualitatives issues des EEE et l'observation d'une restitution organisée par l'un des 4 enseignant(s).

- ***Première étude de cas : un enseignement de français en master 1***

Il s'agit d'un enseignement de français d'une durée de 70 h. en master première année de la spécialité *Enseigner dans le premier degré* du master MEEF. Avec les mathématiques, le français a un poids conséquent dans le master étant donné son importance dans le concours et dans les programmes de l'école primaire. Cet enseignement qui poursuit à la fois l'objectif de préparer au métier (aspects didactiques) et au concours (maîtrise de la langue) vise essentiellement, dans les faits, la préparation du concours qui a lieu en fin de M1 (épreuves écrites) et en fin de M2

(épreuves orales pour les étudiants admissibles à l'écrit). Auparavant, les deux objectifs étaient poursuivis mais l'EEE - que l'enseignante pratique depuis plusieurs années par l'intermédiaire d'un questionnaire ouvert distribué au dernier cours et auquel les étudiants répondent de manière anonyme - avait montré, depuis la mastérisation et l'évolution du concours, une dévaluation de la partie didactique par les étudiants, pour qui cette partie ne devrait pas être abordée en M1 car elle est déconnectée des épreuves écrites du concours. Suite à ces appréciations et également en raison du nouveau contexte créé par le concours dont les épreuves d'admissibilité ne comportent pas de volet professionnel, et qui exige une somme de connaissances sur la langue très importante, l'enseignante a décidé de supprimer cet aspect de l'enseignement, tout en le déplorant.

- *Les appréciations des étudiants sont plutôt convergentes* : Les appréciations des étudiants convergent vers un jugement positif de cet enseignement. Peu de divergences dans l'orientation générale de l'appréciation de l'enseignement apparaissent. En général de 1 à 3 étudiants seulement (sur un groupe de 28) ne partage(nt) pas les avis plutôt ou très positifs exprimés. L'analyse, d'un point de vue quantitatif et qualitatif, des aspects les plus particulièrement appréciés, fait apparaître le profil d'un enseignement qui se structure autour de lignes de force relatives à une orientation claire vers la préparation du concours, un climat favorable à l'apprentissage et une pratique de l'évaluation le soutenant (travail personnel régulièrement sollicité et donnant lieu à des retours réguliers sur les productions des étudiants leur permettant de progresser, une valeur informative de la note « qui ne compte pas mais permet de progresser »). Ils considèrent surtout avoir progressé en grammaire, en rédaction et en commentaire, même s'ils considèrent qu'il est encore trop tôt pour se prononcer à ce sujet. Peu de divergences non plus dans les aspects recueillant les avis plus critiques : Les étudiants considèrent tous n'avoir pas progressé aux niveaux de l'expression orale, de la pensée critique et du travail en équipe. Quelques divergences s'expriment cependant en ce qui concerne les modalités d'évaluation plus ou moins claires selon les étudiants et la participation des étudiants, sollicitée mais peu effective. Certains se jugent « passifs », d'autres considèrent que l'enseignante leur pose des questions mais que celles-ci sont « inductives ou dirigées », qu'ils n'ont pas le temps d'y réfléchir, que souvent l'enseignante répond à leur place. Ils tendent pourtant à considérer que cette méthode permet « d'avancer plus rapidement ».

- *Concordance entre étudiants et enseignant* : La concordance des attentes des étudiants et de l'enseignante, liée à une interprétation similaire du contexte de cet enseignement conduisant les deux acteurs à privilégier la préparation du concours apparaît à la base du jugement positif de cet enseignement. Lors de la discussion relative aux EEE, l'enseignante exprime la satisfaction de

voir ce point de vue partagé. Semble également apparaître une représentation plus ou moins partagée d'un bon enseignement (dans l'optique de préparer le concours) selon laquelle le cours est essentiellement dirigé par l'enseignant, il comporte de nombreux exemples et des exercices donnant lieu à des feedbacks individuels et collectifs permettant aux étudiants de s'entraîner et de progresser.

- *Les apports du débat entre les étudiants et l'enseignant* : C'est lors de la discussion collective que la contradiction entre un souhait d'avoir plus de didactique et un principe de réalité amenant à privilégier le concours s'exprime de manière plus nette : « On aimerait bien avoir plus de didactique, cela nous intéresserait plus mais c'est impossible ». Par ailleurs, une partie de la discussion collective consécutive aux premières EEE fut consacrée à la prise de notes jugée trop importante et peu profitable par la plupart des étudiants considérant que des exercices leur seraient plus utiles ; quelques-uns cependant jugeant au contraire que cette activité de prise de notes leur permettait de fixer des connaissances. Lors de la seconde discussion, deux mois plus tard, les étudiants exprimèrent leur satisfaction que l'enseignante ait pris en compte leurs attentes à ce propos en leur distribuant des supports de cours plus importants. Certains firent part de l'évolution de leur point de vue à ce propos « au départ je faisais partie de ceux qui étaient plutôt favorables à la prise de notes mais j'ai finalement constaté que je progresse plus en écrivant moins et en faisant plus d'exercices ». On voit donc à travers cet exemple que la discussion a permis de faire émerger un élément non prévu dans le questionnaire, relatif aux modalités du travail des étudiants dans les cours et à la grande part consacrée à la prise de notes. Par l'expérimentation « conscientisée » d'une autre pratique donc à une réflexion sur les processus d'enseignement-apprentissage, des points de vue contradictoires ont pu être exprimés et conduire à des modifications de la pratique de l'enseignante mais également à l'évolution des points de vue des étudiants sur leur représentation du travail des étudiants durant les cours.

- *Décalages et régulation* : L'enseignante a découvert dans les EEE des aspects de l'activité des étudiants qui ne vont pas dans le sens de ses objectifs : « Je suis très étonnée par le fait qu'ils ne considèrent pas avoir progressé dans le travail en équipe c'est pour moi un objectif très important et je leur demande de faire le travail personnel à plusieurs mais visiblement ils ne le font pas ». Ce constat décevant l'a amenée à prendre certaines décisions allant dans le sens d'exercer plus de contrôle, un guidage plus contraignant de l'activité des étudiants : « La prochaine fois quand je vais leur donner le travail ils devront le rendre par groupe. Je vais ainsi créer l'obligation de travailler en équipe. » C'est la lecture d'un commentaire isolé, émis par un(e) étudiant(e) « On est parfois un peu infantilisés », qui a produit chez l'enseignante, lors de l'entretien, le début d'un réel questionnement réflexif sur sa pratique. Bien qu'un(e) seul(e)

étudiant(e) l'ait écrit cela a interrogé l'enseignante et c'est d'ailleurs sur cette phrase qu'elle a quitté notre entretien. A quoi cela renvoie-t-il ? L'interprétation de l'enseignante est que : « cela pourrait venir du fait qu'ils sont mis dans la situation de ne pas maîtriser les connaissances qu'ils devraient maîtriser qui renvoient à un passé scolaire lointain. Le fait de constater qu'ils ne font finalement que ce qui est exigé dans le court terme avec rendu, je me demande si ce ne sont pas eux qui s'infantilisent ». D'un autre côté, l'enseignante s'interroge sur sa propension à « materner » les étudiants. On peut observer le début d'un questionnement de certains aspects de cette représentation partagée de l'enseignement, l'enseignante s'interrogeant sur sa pédagogie et les étudiants construisant progressivement une représentation mieux partagée de certains aspects de l'activité des étudiants en cours.

- ***Deuxième étude de cas : un enseignement de physique en Master 1***

L'enseignement de physique, qui est pris dans la même double contrainte de préparer au métier et au concours, dispose d'un volume horaire bien plus réduit. L'objectif de cet enseignement est de travailler sur la culture scientifique des savoirs de base envisagés comme des savoirs à enseigner. Il est fondé sur des mises en situation amenant à mobiliser ces connaissances dans un raisonnement scientifique pendant les cours et des parcours pédagogiques numériques complémentaires à étudier en autonomie en dehors des cours, parcours construits par l'enseignant pour faire face au déficit d'heures de cours qui ne permet pas de revoir toutes les notions requises pour le concours. L'étude personnelle de ces parcours pédagogiques fait partie du contrat défini par l'enseignant. L'enseignant (P1) vise donc à préparer aux attendus du concours tout en construisant des fondamentaux épistémologiques et méthodologiques nécessaires à l'enseignement scientifique, sachant que les aspects didactiques sont traités en 2ème année de master. Il privilégie la construction des savoirs par l'évaluation formative en classe dans des modalités interactives, souvent de travail collectif à partir de productions individuelles ou collectives faites en classe.

- *Convergences et divergences entre les étudiants au sujet de cet enseignement* : De fortes convergences dans les opinions exprimées par les étudiants soulignent les lignes de force de l'enseignement : enthousiasme de l'enseignant, intérêt du cours, utilisation de méthodes actives (et notamment les boitiers de vote qui recueillent une appréciation très positive), bon climat, enseignement appuyé sur de nombreux supports et exemples. Ce sont les items qui recueillent les meilleurs scores et également des commentaires exclusivement positifs. On observe également des convergences fortes dans les appréciations plus critiques relatives à la connaissance des attendus pour la validation du master, à la lourdeur des parcours pédagogiques qu'ils utilisent peu, ainsi qu'au manque de suivi individuel. Les étudiants considèrent dans l'ensemble être plutôt

bien préparés au métier, même si des voix très minoritaires divergent. Pour cet item, sur 18 commentaires exprimés 15 vont dans le même sens et 11 sont quasiment identiques à celui-ci : « des expériences faciles à reproduire en classe ». Seulement 2 expriment qu'« il n'y a pas assez de didactique ». Mais quelques items du questionnaire recueillent des avis contrastés. Il s'agit surtout d'appréciations différentes de la difficulté du cours et du rythme de l'enseignement. Les étudiants ayant un parcours scientifique trouvent qu'ils ont peu de travail personnel à fournir, que le rythme est adapté voire lent, qu'ils apprennent peu de nouvelles notions alors que des étudiants plus nombreux, au profil généralement littéraire, expriment des difficultés à suivre le rythme des cours sans avoir le temps de travailler suffisamment les parcours pédagogiques et trouvent le programme lourd. En ce qui concerne la préparation au concours alors que des étudiants estiment « qu'il y a un lien assez fort » d'autres regrettent au contraire « peu d'entraînements type concours ». De même certains ont déjà réalisé les sujets mis à leur disposition dans les parcours alors que d'autres ne les ont pas encore faits.

- *Convergences et divergences entre enseignant et étudiants* : les convergences se manifestent sur des aspects appréciés de l'enseignement qui conduisent alors l'enseignant à exprimer une satisfaction, un sentiment de reconnaissance professionnelle. L'enseignant peut également converger avec les étudiants sur certaines critiques qui donneront ou non, lieu à une régulation de l'enseignement selon que cet enseignant, P1 estime qu'elles sont ou non sous son contrôle. Ainsi il reconnaît ne pas avoir été suffisamment clair sur les exigences de l'examen et pouvoir facilement y remédier : « C'est vrai qu'on l'a formulé mais peut-être encore trop à l'oral. Ça se peut bien, j'y ferai attention. C'est des choses où on peut les rassurer très facilement en l'écrivant ou en leur donnant des sujets des années précédentes. Ça, ce n'est pas difficile ». En revanche si P1 reconnaît que le besoin exprimé des étudiants d'avoir plus de retours individuels sur leur travail est légitime, il estime n'être pas en mesure d'y répondre pour des raisons d'une part contextuelles (les étudiants sont trop nombreux et il n'a pas assez d'heures avec eux), d'autre part pour des raisons qui tiennent à des limites personnelles « Il reste plein de choses à inventer mais là je touche à mes limites. Sur la gestion de ce genre de trucs, je risque de dérailler, regarder qui fait quoi, faire tourner... je sais que je ne suis pas assez structuré pour ça ». Il estime également que la multipolarité de la formation est un frein majeur. « Cette multipolarité dans laquelle on se trouve englués c'est normal qu'ils la ressentent aussi », de même que le niveau trop hétérogène des étudiants « On est dans des situations impossibles à tenir, à un moment donné il y a un côté insupportable d'avoir des étudiants au niveau master qui n'ont pas le niveau collègue. Il faut le dire à un moment donné. Il faudrait avoir contractualisé le fait que quand ils arrivent à l'IUFM ils maîtrisent déjà les connaissances de base. »

De fortes divergences s'expriment quant à la responsabilité de l'étudiant dans le processus d'enseignement-apprentissage. L'enseignant attend plus d'autonomie et d'investissement dans l'organisation du travail personnel, là où les étudiants attendent un guidage plus serré et individualisé. La position radicale exprimée dans le commentaire suivant dépasse un positionnement individuel circonscrit à l'enseignement de physique. Elle remet en question le modèle de formation général : « Moi je pense qu'à force on tue leur autonomie et on les sclérose. Ils devraient pouvoir travailler seuls et rendre compte des compétences développées à leurs enseignants, parce qu'il faudra forcément qu'ils s'y collent tout seul, ils ne vont pas enseigner la même chose toute leur vie... en plus on ne les normalise même plus comme avant, encore à l'école normale ils apprenaient des façons de faire rigides certes mais au moins ils apprenaient quelque chose. Si l'idée, c'est d'éviter cette normalisation et de développer un niveau d'autonomie qui leur permette de se former et d'acquérir des compétences seuls, alors là on n'y est pas du tout ! On se trompe. Si on ne change pas radicalement les façons de faire et si on ne met pas l'étudiant au centre en exigeant de sa part un travail, on court à notre perte... » (P1). Enfin l'EEE a révélé que la visée formative des modalités d'évaluation mises en œuvre par l'enseignant pouvait ne pas être perçue de la même manière par certains étudiants de par leur vision normative classante de l'évaluation « Je l'ai vécu comme quelque chose de stressant. J'avais peur d'être repérée directement comme mauvaise élève » ; « au début on se sent mal de répondre faux » ; « c'est vrai qu'en début de cours on est déstabilisé parce qu'on ne sait pas répondre aux questions. Ça serait à la fin on pourrait voir ce qu'on a compris ». La discussion a amené l'enseignant à s'interroger également sur la pratique de l'évaluation au sein de l'IUFM : « Ils sont déjà dans une situation de concours et en plus ce qu'on fait dans le master renforce cet individualisme... pratiquement aucune évaluation sommative collective, y en a-t-il une d'ailleurs ? »

On voit donc à travers cette étude de cas, que l'analyse des EEE génère des réflexions touchant aussi bien à la personne de l'enseignant qu'à des aspects techniques et qui débordent largement le champ circonscrit d'un enseignement, pour concerner le collectif pédagogique et ses choix de formation, in fine le modèle d'enseignement/apprentissage construit à travers les dispositifs de formation. Si l'enseignant envisage bien des régulations tenant surtout à une meilleure explicitation du contrat, ses limites personnelles mais aussi les conditions de l'enseignement, (bipolarité de la formation ; place de l'enseignement dans l'ensemble ; modalités d'évaluation) en bornent les possibles.

- ***Troisième étude de cas : un enseignement de français en master 2 par deux enseignants***

La troisième étude de cas concerne deux enseignants assurant l'enseignement de français en master 2, chacun gérant un groupe d'étudiants. L'enseignement de master 2 qui suit celui du M1, est également pris dans la double contrainte de préparer les étudiants à la fois à la réussite du concours et de les former à leur métier. A cette contrainte se rajoute un public composé pour partie d'étudiants ayant réussi l'écrit du concours et se préparant à l'épreuve orale et pour partie d'étudiants ayant échoué en M1 et repassant les épreuves écrites à la fin du M2. Là aussi les deux enseignants déplorent cette modalité qui conduit à une tension entre les objectifs de professionnalisation et de réussite du concours et qui amène les étudiants à se centrer sur l'obtention du concours, pour finalement considérer qu'ils n'ont pas été assez formés au métier. Les deux privilégient cependant la formation des maîtres. La spécificité de cette étude de cas par rapport aux deux précédentes est que la démarche d'évaluation de l'enseignement est portée par trois étudiants, à l'exception de la discussion finale entre les deux enseignants. C'est un élément de contexte important car le processus de discussion se déroule alors dans un cadre orienté par des positionnements catégoriels (étudiants/enseignants), même si le dispositif suit le format général décrit plus haut.

- *De fortes convergences dans les avis des étudiants* : une forte convergence s'exprime entre les étudiants sur de nombreux items. Ils reconnaissent avoir un enseignant enthousiaste, au discours clair et structuré et qui sollicite la participation active des étudiants. Mais de manière unanime, les étudiants jugent que cet enseignement ne prépare pas assez au concours : « le cours ne prépare pas à l'oral », c'est-à-dire « qu'il n'y a pas assez de mises en situation comme le jour du concours ». De même leur jugement négatif sur la préparation au métier s'explique par un contenu jugé trop théorique : « Conseils intéressants mais éloignés de la réalité du terrain » ; « trop peu d'éléments utilisables en classe » ; « de bons éléments théoriques mais pas toujours transférables sur le terrain », « on ne sait pas faire une séquence, paradoxal pour un professeur d'école ». Ils mettent également en évidence la conscience de décalages entre le projet de l'enseignant et celui des étudiants : « l'enseignant trouve sûrement son travail pertinent dans le cadre de la formation de futurs enseignants mais pour les étudiants ça ne l'est pas ». Ces décalages importants conduisent à une désaffection du cours, exprimée par l'absentéisme et une faible participation : « Je ne participe pas car le cours n'est pas en adéquation avec ce que j'attends ». Du fait de ces convergences, la discussion organisée entre les étudiants a apporté peu d'informations supplémentaires et n'a pas donné lieu à débat puisque les étudiants étaient d'accord entre eux. Elle a cependant permis d'élaborer des propositions d'évolution concrètes qui

n'avaient pas été exprimées dans les questionnaires individuels. Pourtant quelques divergences d'appréciation exprimées dans les questionnaires auraient pu servir de base de discussion. Ces divergences apparaissent aux niveaux de l'intérêt du cours, de la préparation au métier et du travail personnel des étudiants.

- *De forts décalages entre enseignants et étudiants* : comme dans les cas précédents, les convergences se manifestent sur des aspects appréciés de l'enseignement, les enseignants apportant des précisions sur leur conception de l'enseignement, et pour l'un d'entre eux sur son ressenti, le second restant plus dans la réserve lors de l'entretien avec les étudiants. Les enseignants peuvent également converger avec les étudiants sur certaines critiques qui donneront ou non, lieu à une régulation de l'enseignement selon qu'ils estiment qu'elles sont ou non sous leur contrôle, ce qui ne semble pas être le cas en ce qui concerne la conciliation du double objectif du concours et du métier ainsi que l'illustre le verbatim suivant : « *J'ai vraiment eu le sentiment que de par la façon dont ont été mis en place les enseignements, les attentes qu'il y avait justement avec ce concours d'un côté la professionnalisation de l'autre nous en tant que formateurs on se retrouvait coincé dedans* ». Il apparaît également que si les enseignants et les étudiants peuvent avoir des appréciations partagées par exemple sur le manque de participation des étudiants, ils ne les interprètent pas nécessairement de la même façon. Alors que les enseignants incriminent la passivité et le manque de maturité des étudiants : « *chloroformés, immatures et manquant de niaque pour le concours, infantiles* », « *On a l'impression de collégiens qui attendent parce que, sur le fait par exemple que l'on n'indique pas nos critères d'évaluation pour le master à l'avance et qu'ils ne savent pas exactement sur quoi ils seront évalués, moi ça me rappelle les élèves de collège* » ; les étudiants mettent en cause l'inadéquation des cours à leurs attentes : « *Non je ne participe pas car le cours n'est pas conforme à ce que j'attends.* » De forts décalages de représentations s'expriment donc à travers le processus d'évaluation, non seulement en ce qui concerne les attentes réciproques ainsi que nous venons de le mettre en évidence, mais aussi dans les objectifs de l'enseignement. La découverte des points de vue des étudiants amène l'un des deux enseignants à déclarer que les étudiants ne peuvent « raccrocher » l'objectif qu'il a annoncé à leurs représentations : « *ni à la façon dont ils se représentent le métier, ni à la façon dont ils se représentent le concours* ». Là où l'enseignant souhaite les faire réfléchir sur les savoirs, les étudiants souhaitent travailler sur la construction de séquences. « *C'est pour nous enfin pour moi intraitable, parce que tu ne peux pas préparer comme ça, pouf, une séance sur les écritures tâtonnées sans avoir un bout de réflexion sur comment ça marche un système d'écriture et qu'est-ce que c'est qu'une écriture* ». Là où l'enseignant souhaite faire un travail de fond, les étudiants attendent des connaissances utiles pour l'action à court terme. « *Mon*

enseignement ne vise pas la rentrée de septembre, il vise ce que sera le maître, trois ans plus tard, quand il aura construit des routines de tenue de classe qui lui permettront justement de retrouver une certaine disponibilité intellectuelle ». La discussion avec les étudiants amène l'enseignant à prendre conscience de l'étendue des décalages. *« J'ai eu le sentiment fondamentalement d'un malentendu sur le métier lui-même.... J'ai eu le sentiment qu'elles se représentaient le métier de façon finalement très applicationniste, comme si c'était clair ce qu'il y a à enseigner et la façon de l'enseigner comme si on était en didactique des maths quoi, en gros d'ailleurs elles le disent à plusieurs reprises il y a souvent une comparaison avec la didactique des maths. Elles pensent qu'il y a des certitudes, elles voudraient les atteindre alors qu'il me semble que justement l'objet d'enseignement, il est un peu obscur en français, que comme il s'agit de travailler la langue, de travailler à travers la parole mais qu'on en est aux balbutiements pour savoir comment est-ce que les enfants acquièrent la parole et acquièrent la langue et acquièrent les compétences de lecture et d'écriture et les compétences langagières, on est loin d'avoir une description des obstacles qui se posent aux élèves et que donc être prof de français c'est quand-même beaucoup dans une réactivité par rapport aux élèves. »*

Comme dans les cas précédents, les résultats quantitatifs et qualitatifs des EEE dessinent une sorte de profil d'enseignement perçu. Cette perception est fonction de l'adéquation du cours aux attentes des étudiants, elles-mêmes fondées sur des représentations du métier et du concours. Les forts décalages entre les attentes des étudiants et les objectifs des enseignants sont à la base d'un jugement plutôt négatif de cet enseignement du côté des étudiants comme du côté des enseignants. Les discussions entre enseignants et étudiants ont permis de prendre la mesure de ces décalages ainsi que l'expriment les enseignants lors de l'entretien collectif des enseignants de français. Ils ne projettent pas d'ajuster leur enseignement aux demandes des étudiants, contrairement à la première étude de cas, et envisagent des régulations d'un autre type que nous présenterons dans la partie suivante.

4.3.4. Types d'effets repérés dans les études de cas

Dans les trois cas présentés, la durée (entre 1h30 et 2h45) et le contenu des entretiens menés sur la base des résultats des EEE montrent l'intérêt porté par les enseignants aux points de vue des étudiants et à leur discussion, principalement focalisée sur les aspects critiqués de l'enseignement. L'analyse des discours des enseignants met en évidence 7 catégories d'effets engageant des dimensions socio-affectivo-cognitives et des dimensions de transformation pédagogique.

- *Des émotions ressenties*

Bien que les enseignants aient tendance à centrer leurs analyses sur les EEE critiques, ils repèrent néanmoins aussi des commentaires montrant que les étudiants (ou certains d'entre eux) ont perçu leur projet pédagogique. Ce qui les amène à exprimer parfois un sentiment de plaisir. « *Ils ont bien compris que l'objectif c'est la préparation au concours et pas au métier.* » « *Quelque part ça me fait plaisir qu'ils perçoivent qu'il s'agit d'un cours ancré dans une forme de pédagogie particulière* ». L'émotion « négative » est peu exprimée mais elle est perceptible à travers la communication non verbale et l'emploi de locutions comme « *je suis déçu* », « *ça me fait bondir* », manifestant déception ou colère.

- *Surprise questionnante*

Certains résultats étonnent les enseignants et les amènent plus ou moins à s'interroger. Il peut s'agir d'une tendance dans les EEE qui permet de découvrir des éléments concernant les étudiants, le peu de travail collaboratif par exemple. Il peut aussi s'agir d'un commentaire isolé, qui fait particulièrement écho et conduit à s'interroger sur son style d'enseignement : « *Ils sont dans une certaine passivité bien installée que peut-être je conforte par le fait de les mater un petit peu et peut-être que je ne leur dis pas qu'ils doivent davantage travailler, peut-être que ce n'est pas suffisamment clair* ». Cet effet de surprise peut conduire ou non à une réflexion approfondie sur les résultats et à des régulations dans l'enseignement. Dans le cas de cette enseignante, on a pu observer l'amorce d'une réflexion qui au départ met en cause l'étudiant pour progressivement en arriver à s'interroger sur sa propre pédagogie et la régulation envisagée qui va dans le sens d'un suivi plus serré de l'activité des étudiants, un choix conforme au style de l'enseignant.

- *Meilleure compréhension des étudiants et de leur perception de l'enseignement*

En tout état de cause, plusieurs commentaires indiquent que les EEE conduisent les enseignants à mieux comprendre leurs étudiants et la perception qu'ils ont de leur enseignement. « *Je l'ai découvert à travers leurs commentaires* ». « *Je trouve ça très difficile d'évaluer mon travail et quand les étudiants eux l'ont évalué parce qu'ils ont fait des pourcentages, j'aurais été incapable moi de donner le moindre indice de satisfaction que pouvaient avoir les étudiants. Parfois il y avait des cours pendant deux heures j'en entendais deux parler. Ça allait mieux au fur et à mesure que l'année avançait mais...* » La discussion entre les enseignants lors de l'entretien collectif a posteriori fait émerger le constat que la faible participation des étudiants en cours n'est pas forcément due à un manque de travail mais à la peur du regard des autres, puisqu'il apparaît qu'ils ne renâclent pas à se mettre au travail quand un écrit leur est demandé mais que c'est la prise de parole en public qui pose problème. Un enseignant relate un incident qui l'a conduit à

faire ce constat quand une étudiante lui a explicitement dit qu'elle avait peur de la façon dont les autres allaient la regarder. Ce rapport à l'évaluation a également été discuté avec l'enseignant de physique lors de l'entretien final, comme nous l'avons analysé plus haut.

- *Mise en évidence des décalages de représentations entre enseignants et étudiants*

La part la plus importante des discussions a été consacrée aux décalages de points de vue entre les enseignants et les étudiants mis en évidence par les EEE qui portent sur les représentations de l'activité de l'étudiant, des objectifs de l'enseignement et du métier de professeur d'école. Plusieurs commentaires des enseignants expriment un point de vue selon lequel les étudiants doivent consentir à plus d'efforts pour s'approprier le cours par un travail personnel plus autonome et engagé, alors que les étudiants semblent plutôt attendre en majorité, un guidage serré de leur activité. Le désengagement est attribué à l'immaturation et à la passivité des étudiants, là où les étudiants, dans un cas évoquent l'inadaptation du contenu du cours qui les conduit à déclarer travailler à partir d'autres sources jugées plus utiles et dans un autre cas, des modalités d'enseignement magistrocentrées (pourtant plébiscitées car jugées en phase avec les attentes). Un enseignant regrette que : « *Les étudiants ne s'autorisent pas à dire ce qu'ils pensent du cours dans le cours et moi je trouve ça particulièrement infantile* » ce qui illustre bien le décalage entre les deux acteurs.

S'expriment également des décalages sur la représentation des objectifs de l'enseignement (concours et métier). Là où les étudiants attendent un enseignement qui les prépare aux épreuves du concours et à l'exercice du métier à travers des tâches et des mises en situation concrètes (épreuves, préparation de séquences), trois enseignants sur quatre privilégient les apports didactiques de fond pensés « *pour le long terme* » et reprochent aux étudiants « *une vision applicationniste du métier* », considérant que leurs représentations les empêchent de juger l'enseignement à sa juste valeur : « *En fait cela vient de leurs représentations. Ils ne savent pas vraiment de ce qu'est le concours. Et donc ils ne savent pas que je les y entraîne* » ; « *Il faut se rappeler que les objectifs du master ce n'est pas que la préparation au concours il faut tenir tous les bouts... Le concours ce n'est pas ma priorité, ils développent des tas de compétences utiles pour le concours mais je serais presque jusqu'au boutiste sur ce point, c'est leur problème le concours* ». Les décalages portent notamment sur le fait qu'ils ne travaillent pas assez sur la construction de séquences : « *C'est pour nous enfin pour moi intraitable, parce que pour pouvoir construire une séquence sur par exemple les écritures tâtonnées il faut pouvoir lire ce que lisent les enfants, ce que font les enfants avec des zones d'incertitude et un questionnement, chez le maître il faut être un peu au clair dans sa [tête] sur ce que c'est qu'un système d'écriture. Et tu ne peux pas préparer*

comme ça, pouf, une séance sur les écritures tâtonnées sans avoir un bout de réflexion sur comment ça marche un système d'écriture et qu'est-ce que c'est qu'une écriture ».

- *Intentions de régulation*

Les quatre enseignants envisagent des régulations différentes en fonction de leur jugement. Une meilleure compréhension des étudiants peut parfois déboucher sur des perspectives d'ajustements, mais les décalages de points de vue et les obstacles perçus peuvent apparaître comme des points de blocage du processus d'évolution de l'enseignement. Les régulations peuvent aller dans le sens des demandes des étudiants quand ils estiment qu'elles sont légitimes et accessibles, en l'occurrence l'explicitation du contrat en précisant les attendus de l'évaluation (cas de l'enseignant de physique) ou une adaptation des modalités d'enseignement (chez l'enseignante de français qui a réduit les moments d'exposés magistraux au profit de plus d'exercices). Il peut également s'agir de modifications non demandées par les étudiants mais qui correspondent à l'analyse que l'enseignant fait des comportements des étudiants non conformes à ses attentes (ex : mise en place de rendus collectifs pour amener les étudiants à travailler en équipe). En revanche les deux enseignants de français de master envisagent des régulations qui ne vont pas nécessairement dans le sens des attendus des étudiants. L'un prévoit de renforcer la confiance en eux des étudiants, leur sentiment de compétence et réfléchit à des solutions pour détendre l'ambiance de la classe. *« Pour moi ce qui peut être fécond, au vu des EEE et des entretiens, c'est de leur faire prendre conscience qu'ils savent beaucoup de choses qu'ils ignorent savoir... Ils ne savent pas qu'ils ont des outils à leur disposition qu'ils peuvent utiliser... Ce serait un de mes objectifs de l'année prochaine c'est-à-dire qu'en classe, on puisse rire aussi ».* Par contre l'autre qui estime avoir fait le même constat que les étudiants, projette de : *« provoquer la crise ».* *« Ce qu'il y a de sûr c'est que l'an prochain, j'évoquais les deux crises celle de décembre et celle de fin janvier février c'est sûr que je vais les provoquer de façon à ce qu'il y ait des choses qui s'éprouvent. En me disant, si ça vient pas, il est essentiel qu'il y ait de leur part pour se former un engagement qui n'était pas là cette année. »* Par ailleurs les quatre enseignants ont projeté des évolutions dans les façons d'évaluer les étudiants : *« Les partiels c'est infantilisant ... et donc l'année prochaine on a décidé qu'il n'y aurait plus de partiels, tout se fera par contrôle continu ; et on va leur proposer les mêmes sujets, pour qu'il y ait une harmonisation, même pour les contrôles continus, l'idée de les faire travailler à deux ».*

- *Obstacles à la régulation*

Certains commentaires expriment un accord avec les critiques des étudiants sans que des ajustements soient considérés possibles. Généralement, les raisons invoquées pour justifier cette impossibilité relèvent de déterminations contextuelles. Les enseignants reconnaissent le besoin

des étudiants d'avoir plus de retours individuels sur leur travail mais ils estiment n'être pas en mesure d'y répondre pour des raisons contextuelles tenant d'une part à l'effectif (les étudiants sont trop nombreux et ils n'ont pas assez d'heures avec eux), d'autre part à la multipolarité de la formation (concours, métier, université). Dans un seul cas, il a été fait état d'une limite relevant de l'individu lui-même : « *Il reste plein de choses à inventer mais là je touche à mes limites. Sur la gestion de ce genre de trucs, je risque de dérailler, regarder qui fait quoi, faire tourner... je sais que je ne suis pas assez structuré pour ça* ».

Par ailleurs les choix pédagogiques du collectif peuvent aussi entrer en contradiction avec ceux de l'enseignant : « *Moi je pense qu'à force on tue leur autonomie et on les sclérose. Ils devraient pouvoir travailler seuls et rendre compte des compétences développées à leurs enseignants, parce qu'il faudra forcément qu'ils s'y collent tout seul, ils ne vont pas enseigner la même chose toute leur vie... en plus on ne les normalise même plus comme avant. Encore à l'école normale, ils apprenaient des façons de faire rigides certes mais au moins ils apprenaient quelque chose. Si l'idée, c'est d'éviter cette normalisation et de développer un niveau d'autonomie qui leur permette de se former et d'acquérir des compétences seuls, alors là on n'y est pas du tout ! On se trompe. Si on ne change pas radicalement les façons de faire et si on ne met pas l'étudiant au centre en exigeant de sa part un travail, on court à notre perte...* »

- *Une démarche d'évaluation de l'enseignement jugée productive*

Lors des entretiens de bilan réalisés à froid un mois et demi après la discussion au sujet des résultats, les quatre enseignants ont systématiquement exprimé une appréciation positive de la démarche d'évaluation de l'enseignement telle que conduite. A contrario est souligné le caractère relativement vain de l'EEE conduite quelques années auparavant dans une démarche institutionnalisée « sans incidence finalement », « lettre morte ». L'expérience est jugée enrichissante et productive : « *Globalement quelque chose de positif dans le sens qu'il s'agit d'une démarche dans laquelle on se place dans une dynamique pour améliorer les choses. Pas du temps perdu, ça m'aide à me repérer personnellement, à continuer à être attentif et développer des démarches pour aller vers du mieux. Sans cela, pas de possible rétroaction sur ce qu'on fait.* » Ils ont particulièrement apprécié la médiation d'un tiers dans le dispositif d'évaluation ainsi que les échanges en amont et en aval. Elle est considérée comme porteuse d'une dynamique d'évolution : « *C'est une expérience très enrichissante, j'ai apprécié de ne pas avoir à conduire moi-même cette évaluation parce que le fait que ce soit toi, ça prenait une autre dimension et je trouvais intéressant de pouvoir lire les commentaires des étudiants en étant dans la distance et le fait de ne pas être impliquée directement et demandeur principal de cette évaluation. C'était pour moi important et ça pouvait donner lieu à des ouvertures nouvelles par rapport à des*

améliorations de contenus de cours ou des manières de procéder dans l'approche des enseignements ; « Le simple fait de parler de ce qu'on fait, de l'échanger, tout le monde n'a pas la chance, expérience de la chose, formuler des questions avec raisonnement... Ce qu'on pense comprendre du monde et ce qui est renvoyé vers les autres, avoir le point de vue des étudiants. » A été également soulignée l'importance de penser l'enseignement comme une co-construction et donc de donner une part active à l'étudiant dans ce processus : *« Au moment de l'entretien, j'avais dit nettement ce que je pensais de comment l'universitarisation de la formation a infantilisé les étudiants, sur le côté malfaisant des cours magistraux sur le côté un peu infantile des partiels dont on à propos de quoi il y a pas de corrigé, il y a pas d'annotation etc.» ; « J'ai trouvé que ce questionnaire était pas mal parce qu'il interrogeait les étudiants à la fois sur l'enseignement qu'ils recevaient et en même temps sur comment eux s'engageaient dans leur préparation de master et de concours, c'est-à-dire il y avait des questions sur le travail personnel. Donc ça à mon avis c'est qu'en lisant le questionnaire ça les amenait à s'interroger sur comment eux s'intégraient dans cette formation. C'est intéressant ça, parce que ça les obligeait à adopter un autre point de vue. »* Il apparaît enfin que les enseignants ont particulièrement apprécié les dynamiques de collectives qu'elles concernent les échanges avec le chercheur, avec les étudiants, ou entre enseignants : *« Tu parlais de coproduction, tu parlais des acteurs de leur formation et j'ai trouvé et ça me tient à cœur, j'ai trouvé qu'ils n'étaient pas tellement acteurs de leur formation et j'ai bien aimé cette rencontre aussi parce qu'elle a permis cette discussion. J'ai senti la levée de certains malentendus. C'était comment dire, c'était gratifiant pour moi »*. Le processus d'évaluation a permis d'amorcer un collectif d'analyse ; *« c'est la première fois qu'on s'en parle de façon un peu calée, un peu officielle un peu pour de vrai mais je trouve que c'est bien de pouvoir mener une réflexion collective sur cette question qui est la question de la relation aux étudiants et de l'ajustement de notre enseignement à leurs attentes et comment faire évoluer leurs attentes et leurs représentations du métier pour qu'ils comprennent mieux ce qu'on essaye de faire. Parce que tu vois, ça ne me remet pas en cause fondamentalement dans mon métier de formateur, ça ne me fait pas bouger dans ma représentation de ce que c'est la formation mais ça m'éclaire sur ce que sont devenus les étudiants »*.

4.3.5. Apports de la recherche et approfondissements requis

Cette expérimentation d'un dispositif coopératif dans la mesure où il a établi la confrontation du point de vue de l'enseignant à ceux des étudiants et que les concordances et les écarts ont été mis en débat entre les enseignants s'est articulée autour d'une méthodologie de coconstruction du sens de l'EEE donnant aux enseignants et aux étudiants un rôle déterminant tout au long du

processus. Nous soulignons les principaux apports de cette recherche et les approfondissements qu'elle requiert.

- ***Engagement réflexif des enseignants et des étudiants***

Alors même que nombre de dispositifs institutionnels se heurtent à une inertie relativement improductive, la qualité de l'engagement des enseignants et des étudiants dans les cas étudiés témoigne une nouvelle fois de l'importance des médiations et des échanges pour un processus réflexif constructif d'évaluation de l'enseignement (Marsh, Fleiner et Thomas, 1975 ; Overall et Marsh, 1979 ; L'Hommedieu, Menges et Brinko, 1990, Smith, 2008). Dans ces conditions particulières, expérimentales et coopératives, les étudiants ont été prolifiques, émettant de nombreux commentaires qui se sont révélés utiles aux enseignants pour une meilleure compréhension des étudiants ainsi que de leur perception de l'enseignement ; et ce dans une perspective dynamique d'évolution de leur enseignement (Bocquillon et al. 2015). Par ailleurs, les enseignants eux-mêmes se sont fortement impliqués tout au long de la démarche réflexive d'évaluation. De plus, l'amorce d'un collectif d'analyse a contribué à renforcer l'ensemble bien qu'il semble également avoir entraîné une dynamique de groupe externalisante : tendance des enseignants à attribuer les problèmes aux étudiants et au contexte. Ce qui nous amène à noter qu'une méthodologie propre à l'accompagnement d'un tel collectif doit être élaborée de manière plus approfondie (Marcel et Murillo, 2014).

- ***Déterminations contextuelles***

Les résultats montrent aussi à quel point il est important de contextualiser l'analyse de l'EEE dans un établissement lui-même inscrit dans un environnement socioculturel plus large. Les caractéristiques du recrutement des enseignants en France opéré par un concours situé au milieu du master forment un environnement particulier contraint qui oriente et restreint les choix pédagogiques des enseignants comme l'engagement des étudiants. Quels que soient les efforts de l'enseignant de sciences, la prégnance du français et des mathématiques dans le concours et dans les programmes scolaires conduisent les équipes pédagogiques et les étudiants à privilégier ces matières, les premières en leur accordant beaucoup plus d'heures dans la maquette et les seconds en y accordant plus d'importance. Ce n'est pas pour autant une tâche plus aisée pour les enseignants de ces disciplines qui se retrouvent du coup beaucoup plus soumis à cette pression du concours. Par ailleurs, dans ce contexte global, une conception normative de l'évaluation s'impose et imprègne également la majorité des étudiants. Ils sont ainsi en attente d'un modèle d'enseignant énonçant et transmettant la norme en termes de savoir et de savoir-faire justes ou faux.

- ***Décalages et malentendus***

Le dispositif a permis de révéler les décalages et malentendus qui sous-tendent les relations entre les enseignants et les étudiants qui autrement seraient restés dans l'ombre, ce qui a largement été référencé dans les parties précédentes. On a pu constater l'importance accordée par les étudiants aux manières d'enseigner faisant écho aux dimensions constitutives de la qualité d'un enseignement communément admises dans la littérature : qualité de la structuration, de la communication, des interactions, explicitation de l'évaluation. Les enseignants sont plus ou moins en phase avec ces attentes et certains ne sont pas disposés à évoluer sur ce qu'ils considèrent comme fondamental. Ils attendent un engagement intellectuel et une autonomie des étudiants là où ceux-ci (en majorité) souhaitent des réponses pragmatiques et un accompagnement cadré. En règle générale, les enseignants mettent l'accent essentiellement et quasi-exclusivement sur des facteurs contextuels et sur les attitudes des étudiants inadaptées à la formation qui leur est dispensée (immaturité, passivité...), alors que les étudiants ciblent principalement une manière d'enseigner appropriée à un contexte.

- ***Co construction***

Il apparaît que la dynamique d'intersubjectivité s'est effectuée à différents niveaux emboîtés : la prise en compte du point de vue des étudiants, le dialogue entre différentes perspectives des enseignants et des étudiants, enfin également les discussions dans l'équipe pédagogique. C'est bien par ce travail d'explicitation de l'expérience à partir de son appréciation que des valeurs différentes ont pu être manifestées et objectivées. A partir de là, elles ont pu être parfois mises en débat, condition de leur co construction, dans des formes dialogiques directes et indirectes. Si cette dynamique d'intersubjectivité ne va pas de soi, les collectifs pouvant également exacerber des positionnements catégoriels, et si elle ne conduit pas nécessairement à la résolution de tous les décalages et de toutes les tensions, elle a participé tout du moins à une meilleure compréhension des perspectives des uns et des autres et à l'amorce de transformations individuelles et collectives. Cette approche coopérative et dialogique pourrait être renforcée par une méthodologie suscitant une discussion plus approfondie des différentes perspectives dans l'optique de susciter des échanges plus approfondis entre enseignants et étudiants.

Cette expérimentation d'un dispositif coopératif visant à appréhender les effets de l'EEE quand les enseignants sont impliqués dans la construction du processus confirme son fort potentiel comme révélateur, comme analyseur et comme activateur. Elle est révélatrice du fort impact des représentations (enseignants/étudiants les uns vis-à-vis des autres et vis-à-vis du contexte) dans la dynamique pédagogique et du frein que représentent les malentendus qui peuvent s'ensuivre. Du point de vue de l'activation d'un processus réflexif et adaptatif, il apparaît que sous ces

conditions de mise en œuvre spécifiques, à savoir la validation par les enseignants des questionnaires d'EEE et la mise en œuvre d'une médiation favorisant les échanges et le partage avec une démarche réflexive continue, une amorce d'intentionnalité de changement s'exprime tout en les conduisant à réaffirmer les bases sur lesquelles leur enseignement se fonde. Ces changements peuvent être des modifications dans leurs actions pédagogiques sans remettre en question les fondements sur lesquels leur enseignement se déploie mais au contraire les renforcer en les explicitant. Les dynamiques subjectives et intersubjectives suscitées par un tel dispositif coopératif constituent des conditions propices à une appropriation de l'EEE. Le portage par un chercheur/médiateur dans les études de cas 1 et 2 et par des étudiants dans l'étude 3 a probablement impacté le processus en activant de manière différenciée les positionnements catégoriels des protagonistes.

L'EEE apparaît ainsi prise dans des nœuds de paramètres et des entrelacs de sens. Ses effets ne sont pas mécaniques, ils s'inscrivent dans un processus situé, critique, interprétatif et intersubjectif. Les études de cas révèlent le potentiel régulateur et formatif/transformatif réel de l'EEE qui génère de l'engagement et de la coconstruction introduisant à l'enjeu d'un véritable espace de réflexion et de production collectives et individuelles sur la pédagogie universitaire. Elles mettent également en évidence les limites de l'EEE qui est contrainte dans une zone proximale de régulation. La densité de ses effets est dépendante de la façon dont les trois univers (de l'enseignant, de l'étudiant, du milieu) entrent en interaction et en coopération. Ce dispositif coopératif réflexif, construit dans le cadre de cette expérimentation, est à considérer comme médiation psycho-socio-technique. Bien que limité, ce dispositif synergique apparaît suffisamment significatif et productif pour envisager la pertinence de sa reprise et de son élargissement. De tels processus d'intersubjectivation engagent fortement la dimension symbolique de la relation éducative qui passe par un jeu du pouvoir et de l'autorité auquel aucun dispositif coopératif ou participatif n'échappe. L'EEE mettant en question les équilibres habituels, les frontières entre les parties prenantes sont à redéfinir.

Conclusion

L'étude théorique, critique et opératoire des pratiques et dispositifs d'évaluation dans l'enseignement et des effets de transformation nous a conduite à nous situer par rapport à la diversité des approches de recherche et des paradigmes engagés. Au fil du temps, la perspective d'écologies des pratiques d'évaluation dans l'enseignement ainsi que la mise à l'étude de l'hypothèse du paradigme d'évaluation synergique dans des dispositifs psycho-socio-techniques sont devenues les lignes directrices de mes recherches.

L'évaluation étant polysémique, plus d'une façon d'évaluer coexistent et nombre de dissonances s'expriment mettant en évidence l'importance de prendre la mesure des logiques paradoxales et contradictoires des milieux d'évaluation. Tout l'enjeu du chantier des synergies n'est pas la réalisation de synthèses abstraites qui résoudraient magiquement les divergences et les conflits en cherchant à les faire disparaître mais bien au contraire vise à les affronter et à les faire tenir ensemble. Au fil du temps, il peut être observé en quoi et comment les dispositifs d'évaluation, situés, coopératifs et appropriés, sont pensés moins comme procédure que comme dynamique de champs relationnels, capables potentiellement de créer des conditions propices à des déploiements créatifs entre parties prenantes. Cette approche fait écho à l'ampleur qu'a pris le coopératif dans le monde de l'évaluation comme dans bien d'autres domaines. Ainsi, dans le champ de l'enseignement universitaire, il est de plus en plus question de production de collectifs incluant collègues et étudiants, de communautés de pratiques.

Pourtant, nos travaux comme tant d'autres montrent que cette culture du coopératif est loin d'être partagée, soulignant l'importance cruciale de mise en place de dispositifs appropriés pour établir une culture du collectif pédagogique. En l'absence de tels dispositifs pour s'accorder ou entrer en résonance sur les pratiques du métier tout en assumant les différences de points de vue et les conflits de critères, c'est le poids institutionnel et les dynamiques relationnelles aléatoires qui en viennent à supplanter les ajustements mutuels.

Comme nous l'avons détaillé dans ce mémoire d'HDR, plusieurs auteurs ont proposé des cartographies des modèles ou paradigmes d'évaluation, tels De Ketele (1993, 2016) ; Guba et Lincoln (1989, 1994) ; Mertens (2015) ; Shadish, Cook et Leviton, (1991), Vial

(2012). L'analyse de ces différentes cartographies, de leurs empiètements et leurs contradictions, nous a amenée à élaborer une autre cartographie des différentes générations d'évaluation qui tout en reprenant les précédentes, repère de nouveaux contours.

Le chantier de la première génération d'évaluations qui a largement porté les efforts sur l'élaboration d'outils de mesure, d'explicitation de référentiels et d'indicateurs de performance individuelle et les tests standardisés se renforce en liaison avec les outils numériques et avec des méthodologies consolidées. Une deuxième génération prenant la mesure des limites du paradigme de la mesure a souligné l'importance d'associer mesure et régulation, quantitatif et qualitatif, standardisation et contextualisation, afin de disposer d'éléments objectivables situés. Une troisième génération, a mis l'accent sur l'importance de la construction du sens individuel et partagé. Une quatrième génération vise à faire tenir ensemble les différentes dimensions de la culture de l'évaluation dans une dynamique synergique de milieu.

Cette réflexion épistémologique sur les paradigmes et les différentes générations qui coexistent m'a conduite à mettre en relation des univers théoriques et pratiques pluriels. Afin de n'évacuer ni l'incertain, ni les limites de la connaissance, ni les paradoxes, et de se prémunir des simplifications, des systèmes clos et d'une orientation dualiste, il est visé de mobiliser et interroger les différentes strates engageant à penser, construire, expérimenter l'ouverture de possibles. Ceci amène à cibler les interactions et inter-engendremments impliqués et à prospecter plus particulièrement dans quelle mesure et comment les dispositifs d'évaluation et les transformations peuvent s'articuler de manière féconde : notamment à savoir comment, les dispositifs d'évaluation normalisante et de contrôle peuvent s'articuler avec des dispositifs d'évaluation-développement formatif, parce que c'est dans l'« entre » de différentes évaluations que les synergies sont recherchées.

Tout l'enjeu du chantier des synergies n'est pas la réalisation de synthèses abstraites qui résoudraient magiquement les divergences et les conflits en cherchant à les faire disparaître mais bien au contraire celui de viser à les affronter et à les faire tenir ensemble. Au fil des recherches présentées, il peut être observé en quoi et comment les dispositifs d'évaluation, situés, coopératifs et appropriés, pourraient être pensés moins

comme procédure que comme dynamique de champs relationnels, capables potentiellement de créer des conditions propices non seulement à des déploiements créatifs entre parties prenantes, mais aussi à une activation des milieux eux-mêmes.

Trois formes de synergies qui entrent en consonance ou dissonance constituent des lignes directrices à faire tenir ensemble : synergies disciplinaires et interdisciplinaires qui renvoient plus spécifiquement aux références et méthodes mixtes mobilisées, synergies entre parties prenantes qui concernent plus particulièrement les relations dialogiques, synergies écologiques qui s'envisagent en termes de milieu, entité écosystémique s'avérant déterminante dans les dynamiques de transformation, co-évolution et adaptation de la culture évaluative.

Une telle démarche d'évaluation écologique dans l'enseignement ne se veut ni systématique ni dogmatique, mais tend à intégrer les singularités et les multiplicités, les devenirs et le participatif. Les dispositifs d'évaluation étant des espace-temps critiques, que ce soit en termes d'expérience, d'expérimentation et de recherche, le paradigme synergique inscrit le processus d'évaluation dans un ensemble situé, ouvert aux différentes échelles de temps et interdépendances individuelles, sociales, économiques, politiques et culturelles, qui constituent les conditions de l'action.

Références¹²

- Abernot, Y., Gangloff-Ziegler, C. et Weisser, M. (2012). Contribution à l'épistémologie de l'évaluation des enseignements par les étudiants. *Éducation et socialisation*, 32. [En ligne] <http://edso.revues.org/361>, consulté le 8 août 2014.
- Abma, T. A. , et Stake, R. E. (2001). Responsive evaluation: Roots and evolution. In J. C. Greene et T. A. Abma (Eds.), *Responsive evaluation: New directions for evaluation*, No. 92 (p. 7-22). Jossey-Bass.
- Abrahamson, L. (2006). A brief history of networked classrooms: Effects, cases, pedagogy, and implications. In D. A. Banks (Ed.), *Audience response systems in higher education* (pp. 1–25). Information Science Publishing.
- Abrami, P. C., d'Apollonia, S. et Rosenfield, S. (2007). The dimensionality of student Ratings of Instruction : An update on what we know, do not know, and need to do. Dans J. Perry et Smart, J. (dir.), *The Scholarship of Teaching and Learning in Higher Education : An Evidence-based Perspective* (p. 446-456). Springer.
- Adler, E. et Haas, P. M. (1992). Conclusion: Epistemic Communities, World Order, and the Creation of a Reflective Research Program. *International Organization*, 46/1, 367-390.
- Agamben, G. (2007). *Qu'est-ce qu'un dispositif ?* (Traduit par M. Rueff). Rivages.
- Albero, B. (2010a). Une approche sociotechnique des environnements de formation, *Éducation et didactique*, 4 (1), 7-14. <https://doi.org/10.4000/educationdidactique.715>.
- Albero, B. (2010b). La formation en tant que dispositif : du terme au concept, dans B.Charlier et F.Henri (dir.), *La technologie de l'éducation : recherches, pratiques et perspectives*. PUF, coll. Apprendre.
- Albero, B. (2014). La pédagogie à l'université entre numérisation et massification. Apports et risques d'une mutation. Dans G. Lameul et C. Loisy (dir.), *La pédagogie universitaire à l'heure du numérique : questionnements et éclairages de la recherche* (p. 27-53). De Boeck.
- Alcoff, L. and Potter E. (eds) (1993). *Feminist Epistemologies*. Routledge.
- Aleamoni, M. L. (1999). Student Rating Myths Versus Research Facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13 (2), 153-166.
- Allal, L. et Laveault, D. (2009). Assessment for learning/Evaluation soutien d'apprentissage. *Mesure et Evaluation en Education*, 32, 99-106.
- Allal, L. et Mottier Lopez, L. (2005). L'évaluation formative: Revue de publications en langue française. In *L'évaluation formative, pour un meilleur apprentissage dans les classes secondaires* (p. 265-290). Paris: OCDE

¹² Selon les normes APA 7eme édition

- Altet, M (1996). Les dispositifs d'analyse des pratiques pédagogiques en formation d'enseignants: une démarche d'articulation pratique-théorie-pratique, dans C. Blanchard-Laville et D. Fablet (dir.), *L'analyse des pratiques professionnelles* (p. 11-26). L'Harmattan.
- Altet, M. (2016). Gaston Mialaret, une vie et une œuvre entre pensée, action, recherche pour améliorer l'éducation. *Les Sciences de l'éducation - Pour l'Ère nouvelle*, 49(3), 21-38. doi:10.3917/lse.493.0021.
- Anadon, M. (2006). La recherche dite « qualitative » : de la dynamique de son évolution aux acquis indéniables et aux questionnements présents. *Recherches Qualitatives*, 26,1, 5-31.
- Anderson, H. M., Cain, J., et Bird, E. (2005). online student course evaluations: review of literature and a pilot study. *American Journal of Pharmaceutical Education*, 69, 1-5. doi:10.5688/aj690105.
- Annot, E. (2012). *La réussite à l'université : du tutorat au plan licence*. De Boeck.
- Annot, E. (2014). Postface. Dans : Geneviève Lameul éd., *La pédagogie universitaire à l'heure du numérique: Questionnement et éclairage de la recherche* (p. 221-226). De Boeck Supérieur. doi:10.3917/dbu.lameul.2014.01.0221.
- Arduino, J. et Berger, G. (1989). *D'une évaluation en miettes à une évaluation en actes : le cas des universités*. ANDSHA-Matrice.
- Arendt, H. (1978). *La vie de l'esprit : La pensée, vol. 1*, trad. par L. Lotringer, PUF 3e éd., 1992.
- Arthur, L. (2009). From performativity to professionalism : Lecturers' responses to student feedback. *Teaching in Higher Education*, 14, 441-454.
- Aubert, I. (2008). Sujet et intersubjectivité. *Trajectoires* [Online], 2 | 2008, Online erschienen am: 16 Dezember 2009, besucht am 29 Mai 2020. URL : <http://journals.openedition.org/trajectoires/209>
- Avery, R. J., Keith Bryant, W., Mathios, A., Kang, H., et Bell, D. (2006). Electronic course Evaluations: does an online delivery system influence student evaluations? *Journal of Economic Education*, 37(1), 21-37.
- Banaji, M. R., et Prentice, D. A. (1994). The self in social contexts. *Annual Review of Psychology*, 45, 297-332. <https://doi.org/10.1146/annurev.ps.45.020194.001501>
- Barbier, J.M. (1983). Pour une histoire et une sociologie des pratiques d'évaluation en formation. *Revue française de pédagogie*, 63, 47-60. doi : <https://doi.org/10.3406/rfp.1983.1891>
- Barras, H. (2017). Impact émotionnel de l'évaluation de l'enseignement par les étudiants chez les enseignants d'une Haute Ecole en Suisse. *Education et Formation*, e-307, 74-90.
- Beauvois J.-L. (2001). Psychologie sociale, psychosociologie et psychologie sociale appliquée. In J.-L. Beauvois et J.-M. Monteil (Dir.), *La psychologie sociale, Tome 5 : Des compétences pour l'application* (p. 9-21). Presses Universitaires de Grenoble.

- Becquet, V. et Musselin, C. (2004). *Variations autour du travail des universitaires*. http://www.cso.edu/upload/pdf_actualites/variation%20autour%20du%20travail%20des%20universitaires%20rapport%20final.pdf
- Bedin, V. (2009)(dir.). *L'évaluation à l'université. Évaluer ou conseiller ?* Presses universitaires de Rennes.
- Bègue, L. et Desrichard, O. (2013). *Traité de psychologie sociale. La science des interactions humaines*. De Boeck.
- Beran, T. N. et Rokosh, J. L. (2009). The consequential validity of student ratings : What do instructors really think ? *Alberta Journal of Educational Research*, 55 (4), 497-511.
- Beran, T., et Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment and Evaluation in Higher Education*, 30, 593-601.
- Bernard, H. (1992). *Processus d'évaluation de l'enseignement supérieur. Théorie et pratique*. Études Vivantes.
- Bernard, H. (2011). *Comment évaluer, améliorer, valoriser l'enseignement supérieur ?* De Boeck.
- Bernard, H. et Normand, S. (1998). L'évaluation de l'enseignement vécue par des professeurs de disciplines différentes. *Mesure et évaluation en éducation*, 21 (2), 117-134.
- Bernard, H., Postiaux, N. et Salcin, A. (2000). Les paradoxes de l'évaluation de l'enseignement universitaire. *Revue des sciences de l'éducation*, 26 (3), 625-650.
- Berthelot, J. (1991). Dualisme et pluralisme en sociologie. *BMS: Bulletin of Sociological Methodology / Bulletin De Méthodologie Sociologique*, 31, 61-66. <http://www.jstor.org/stable/24311082>
- Berthiaume, D., Lanarès, J., Jacqmot, C., Winer, L. et Rochat, J.-M. (2011). L'évaluation des enseignements par les étudiants (EEE), *Recherche et formation*, 67, 53-72. [En ligne] <https://journals.openedition.org/rechercheformation/1387>.
- Bertrand C., (2014). *Soutenir la transformation pédagogique dans l'enseignement supérieur*. Rapport à la demande de Madame S. Bonnafous, Directrice générale pour l'enseignement supérieur et l'insertion professionnelle. Ministère de l'enseignement supérieur et de la recherche.
- Boice, R. (1992). *The new faculty member: Supporting and fostering professional development*. Jossey-Bass.
- Bourdieu, P. (1976). Le champ scientifique. *Actes de la recherche en sciences sociales*. 2, 2-3, juin La production de l'idéologie dominante. pp. 88-104. doi : <https://doi.org/10.3406/arss.1976.3454>.
- Bourdieu, P. et Wacquant L. (1998). Sur les ruses de la raison impérialiste. *Actes de la recherche en sciences sociales*, 121-122, 109-118. doi : <https://doi.org/10.3406/arss.1998.3250>.

- Brady, P. (1994). How likeability and effectiveness ratings of college professors by their students are affected by course demands and professors' attitudes. *Psychological Reports*, 74, 3, 197-204.
- Brassac, C. (2005). La réception de George Herbert Mead en psychologie sociale francophone : réflexions sur un paradoxe. *Les Cahiers Internationaux de Psychologie Sociale*, 2, 66, 3-14. URL : <http://www.cairn.info/revue-les-cahiers-internationaux-de-psychologie-sociale-2005-2-page-3.htm> DOI : 10.3917/cips.066.0003
- Brewer, C. A. (2004). Near real-time assessment of student learning and understanding in biology courses. *BioScience*, 54(11), 1034–1039.
- Brookfield, S. (1995). *Becoming a Critically Reflective Teacher*. Jossey-Bass.
- Broussal, D. (2017). L'émancipation : approches conceptuelles. Dans J.F. Marcel et D. Broussal (Dir.), *Émancipation et recherche en éducation. Conditions de la rencontre entre science et militance*. Éditions du Croquant.
- Brown, J.D. (2010). High self-esteem buffers negative feedback: Once more with feeling, *Cognition and Emotion*, 24(8), 1389-1404. DOI: 10.1080/02699930903504405.
- Brown, N. (2006). Shifting Tenses: from “Regimes of Truth” to “Regimes of Hope”, *SATSU Working Paper*, 30, 1-17.
- Burden, P. (2008). Does the use of end of semester evaluation forms represent teachers' views of teaching in a tertiary education context in Japan? *Teaching and Teacher Education*, 24, 1463–1475.
- Burgel, G. (2006). *Université, une misère française*. Hachette Littératures, coll. « Tapage ».
- Burns, J.M. (1978) *Leadership*. Harper et Row.
- Caldwell, J.E. (2007). Clickers in the Large Classroom: Current Research and Best Practice Tips. *CBE-Life Sci Educ*, 6(1), 9–20. doi:10.1187/cbe.06-12-0205.
- Campbell, D.T. (1969). Reforms as experiments. *American Psychologist*, 24, 409-429.
- Campbell, D.T. (1984). Can we be scientific in applied social science ? In R.F Connor, D.G. Altman and C. Jackson (Eds.), *Evaluation studies review annual* (vol. 9, pp. 26-48). Sage.
- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, 59(3), 538–549. <https://doi.org/10.1037/0022-3514.59.3.538>.
- Cashin, W.E. (1996). Developing an effective faculty evaluation system. *Idea Paper*, n° 33, Center for Faculty Evaluation and Development, Kansas State University.
- Castoriadis, C. (1975). *L'institution imaginaire de la société*. Seuil.
- Centra, J. A. (1993). *Reflective Faculty Evaluation*. Jossey-Bass Publishers.

- Centra, J.A. (1973). The student as godfather ? The impact of student ratings on academia, *Educational researcher*, 2, 10, 4-8.
- Chan, C.K.Y. Luk, L.Y.Y. and Zeng, M. (2014). Teachers' perceptions of student evaluations of teaching, *Educational Research and Evaluation*, 20:4, 275-289, DOI: 10.1080/13803611.2014.932698.
- Charlier, J.-E., Croché, S., Fallon, C. et Leclerc, B. (2014). En guise de conclusion : sommes-nous arrivés à organiser un débat de qualité entre partisans et analystes critiques des dispositifs qualité dans l'enseignement supérieur ? Dans C. Fallon et B. Leclercq (Eds.) *Leurres de la qualité dans l'enseignement supérieur ? Variations internationales sur un thème ambigu* (p.339-354). Academia-L'Harmattan.
- Charlot, B. (2008). La recherche en Education entre savoirs, politiques et pratiques : spécificité et défis d'un champ de savoir. *Recherches et éducations*, 1, 155-174. [En ligne] : <http://journals.openedition.org/rechercheseducations/455>
- Clark, B. R. (1987). *The Academic Life, small worlds, different worlds*. Carnegie Foundation for the Advancement of Teaching.
- Clot, Y. (1999). *La fonction psychologique du travail*. Presses universitaires de France.
- Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction: a meta-analysis of multisection validity studies. *Research in Higher Education*, 13, 4, 321-341
- Cohen, P.A., and Benson, BA. (1988). Workload and student course ratings in dental school. *Journal of Dental Education*, 52(2):98-101.
- Cole, A. L. (1989). Researcher and teachers: Partners in *theory building*. *Journal of Education for Teaching*, 15, 225-237.
- Cole, A. L., et Knowles, J. G. (1993). Teacher development partnership research: A focus on methods and issues. *American Educational Research Journal*, 30, 473-495.
- Cook, T.D. and Campbell, D.T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin.
- Cosnefroy, L. (2011). *L'apprentissage autorégulé, entre cognition et motivation*. PUG.
- Costin, F., Greenough, W.T. et Menges, R. J. (1971). Student Ratings of College Teaching: Reliability, Validity, and Usefulness. *Review of Educational Research*, 41, 5, 511-535.
- Coulon, A. (1993). *Ethnométhodologie et éducation*. Presses universitaires de France.
- Cranton, P. (2001). Interpretative and Critical Evaluation, *New Directions for Teaching and Learning*, 88, 87- 97. Jossey-Bass.
- Crawford, L. E. D. et Shutler, P. (1999). Total quality management in education : Problems and issues for the classroom. *The International Journal of Educational Management*, 67-72.
- Cross, K.P. et Steadman, M. H. (1996). *Classroom research : Implementing the scholarship of teaching*. Jossey-Bass.

- Crossgrove, K., et Curran, K. (2008). Using clickers in non-majors and majors-level biology courses: Student opinion, learning, and long-term retention of course material. *Life Sciences Education*, 7, 146-154.
- Dahler-Larsen, P. (2005). *Den rituelle refleksion – om evaluering i organisationer*. Syddansk Universitetsforlag.
- Dangel, H. L., et Wang, C. X. (2008). Student response systems in higher education: Moving beyond linear teaching and surface learning. *Journal of Educational Technology Development and Exchange*, 1(1), 93-104. <http://www.sicet.org/journals/jetde/jetde08/paper08.pdf>.
- De Hosson, C., Décamp, N., Morand, E et Robert, A. (2017). Approcher l'identité professionnelle d'enseignants universitaires de physique : un levier pour initier des changements de pratiques pédagogiques. *Recherches en didactique des sciences et des technologies*, 11. [En ligne] <http://journals.openedition.org/rdst/1014> doi:10.4000/rdst.1014
- De Ketele, J.-M. (1993). L'évaluation conjugée en paradigmes. *Revue française de pédagogie*, 103, 59-80. DOI : <https://doi.org/10.3406/rfp.1993.1298>
- De Ketele, J.M. (2016). L'évaluation et ses nouvelles tendances, sources de dilemmes. *Education permanente*, 208, 19-32.
- DeBourgh, G. A. (2008). Use of classroom "clickers" to promote acquisition of advanced reasoning skills. *Nurse Education in Practice*, 8, 76-87.
- Dejean, J. (2002). *L'évaluation de l'enseignement dans les universités françaises*. Haut Conseil de l'évaluation de l'école.
- Depover, C. et Strebelle, A. (1997). Un modèle et une stratégie d'intervention en matière d'intégration des TIC dans le processus éducatif, in L.-O. Pochon, et A. Blanchet (Eds.), *L'ordinateur à l'école : de l'introduction à l'intégration*, (p. 73-98). IRDP.
- Desgagné, S. (1997). Le concept de recherche collaborative: l'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants. *Revue des sciences de l'éducation*, 23, 2, 371-393.
- Desgagné, S. (2007). Le défi de coproduction de savoir en recherche collaborative. Autour d'une démarche de reconstruction et d'analyse de récits de pratique enseignante, in M. Anadon (Ed.), *La recherche participative: multiples regards*, (pp.89-121), Presses de l'Université du Québec.
- Desgagné, S., et Berdnarz, N. (2005). Médiation entre recherche et pratique en éducation: faire de la recherche «avec» plutôt que «sur» les praticiens. *Revue des sciences de l'éducation*, 31,2, 245-258.
- Detroz P. (2014). *Vers une démarche balisant la mise en place (ou la refonte) d'un dispositif d'évaluation des enseignements*. Communication présentée au 28^e Congrès de l'Association internationale de pédagogie universitaire (AIPU), Mons, 19-22 mai 2014.

- Detroz, P. (2008). L'évaluation des enseignements par les étudiants : état de la recherche et perspectives. *Revue française de pédagogie*, 165, 117-135.
- Detroz, P. (2010). *Le point sur les dispositifs d'évaluation des enseignements par les étudiants : pertinence, utilisation, amélioration*. Thèse de doctorat non publiée, Université de Liège.
- Detroz, P. et Verpoorten, D. (2017). De l'évaluation des enseignements à la régulation des pratiques des enseignants : quels possibles et quelles conditions ? *Formation et recherche*, 307 (1). [En ligne] <http://hdl.handle.net/2268/206415>.
- Detroz, P. et Younès, N. (2014). L'évaluation formative interactive avec les boitiers de vote : un cadre intégrateur. Communication présentée au 26e colloque international de l'ADMEE-Europe, Marrakech, Maroc. Récupéré de l'archive HAL : <http://hal.archives-ouvertes.fr>
- Dewey, J. (1993). *Logique. La théorie de l'enquête* (traduit par G. Deledalle ; 2^{ème} éd., paru en 1967). PUF.
- Dewey, J. (2010). *Le public et ses problèmes* (traduit par J. Zask). Gallimard.
- Dewey, J. (2011). *La formation des valeurs* (traduit et présenté par Alexandra Bidet, Louis Quéré et Gérôme Truc). Empêcheurs de penser en rond.
- Dionne, E. (2012). Les systèmes à réponses personnalisées (SRP) : un atout pour faire de l'évaluation formative en salle de classe ? *Mesure et évaluation en éducation*, 35,1, 47-65.
- Doyle, K. O. Jr. (1983). *Evaluating Teaching*. Lexington Books, D.C.
- Draper, S. W., et Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, 20, 81-94.
- Dumont, A, Rochat, J.-M., Berthiaume, D. et Lanarès, J. (2012). Les effets de l'EEE sur le développement professionnel des enseignants et l'expérience d'apprentissage des étudiants. Comparaison de deux cas suisses. *Mesure et évaluation en éducation*, 35 (3), 85-115.
- Duykaerts, C. et Malmedy, M. (2014). Démarches qualité et évaluation externe de la qualité : mais que vise donc exactement l'Agence pour l'évaluation de la qualité de l'enseignement supérieur ? Dans C. Fallon et B. Leclercq (Eds.) *Leurres de la qualité dans l'enseignement supérieur ? Variations internationales sur un thème ambigu* (pp.141-153). Academia-L'Harmattan.
- Earl, L. (2003). *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning*. Corwin Press.
- Edström, K. (2009). Doing course evaluation as if learning matters most. *Higher Education Research and Development*, 27, 2, 95-106. DOI: 10.1080/07294360701805234
- Endrizzi, L. (2017). Recherche ou enseignement : faut-il choisir. *Dossier de veille de l'IFÉ*, 116, École normale supérieure de Lyon.
- ENQA, (2006). *Références et lignes directrices pour le management de la qualité dans l'espace européen de l'enseignement supérieur* (CNE, Trad.). CNE. (Œuvre originale publiée en 2005).

- Entwistle, N. J. (1998). Improving Teaching Through Research on Student Learning. Dans J. F. J. Forest (dir.). *University Teaching. International Perspectives* (p. 73-112). Garland Publishing.
- Ernst, D. (2014). *Un dispositif d'évaluation des enseignements par les étudiants au service de l'amélioration des programmes*. Communication présentée au 26e colloque de l'ADMEE Europe, 15-16-17 janvier 2014, Marrakech.
- Espéret E. (2001). *Nouvelle définition des tâches des enseignants et des enseignants-chercheurs dans l'enseignement supérieur français*. Rapport. Paris : MEN.
- Étienne, R., Annot, E., Biaudet, P. (2018). Les enseignants-chercheurs débutants en France : l'urgence de la pédagogie. In D. Adé et T. Piot. *La formation entre universitarisation et professionnalisation - Tensions et perspectives dans des métiers de l'interaction humaine*, (p. 67-89). PURH.
- Feldman, J. (2002). Objectivité et subjectivité en science. Quelques aperçus. *Revue européenne des sciences sociales*, XL-124, 85-130. URL : <http://journals.openedition.org/ress/577> ; DOI : 10.4000/ress.577
- Feldman, K. A. (1976). The Superior College Teacher from the Students' View. *Research in Higher Education*, 5, 243-288.
- Feldman, K.A. (2007). Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings¹. In: Perry R.P., Smart J.C. (eds) *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*. Springer.
- Fessler, R. et Christensen, J.-C. (1992). *The teacher career cycle: Understanding and guiding professional development of teacher*. Allyn and Bacon.
- Figari, G. (1994). *Évaluer, Quel référentiel ?* De Boeck.
- Figari, G. (2001). Us et abus de la notion de référentiel. Dans G. Figari et M. Achouche (dir.), *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels* (p. 310-314). De Boeck.
- Figari, G. et Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation*. De Boeck.
- Finkelstein, M. J. (1984). *The American Academic Profession : A Synthesis of Social Scientific Inquiry since World War II*. Ohio State University Press.
- Fischbach, F. (2000). Reconnaissance et philosophie pratique chez Fichte. In M. Bienenstock et M. Crampe-Casnabet (dir.), *Dans quelle mesure la philosophie est pratique ? Fichte, Hegel*, (p. 131-146). ENS Éditions.
- Fitzpatrick, J.L., Sanders, J.R., and Worthen, B.R. (2011). *Program evaluation: Alternative approaches and practical guidelines*. (4th ed.). Pearson.
- Fontan, J.M. et Lachance, E (2005). Pour une évaluation de cinquième génération. *Cahiers de l'ARUC-ÉS*, n°18. ISBN 2-89276-389-4
http://www.perfeval.pol.ulaval.ca/sites/perfeval.pol.ulaval.ca/files/publication_196.pdf

- Forest, L. (2009). Une approche intégrée d'évaluation, d'amélioration et de valorisation de l'enseignement dans une école polytechnique. Dans M. Romainville et C. Coggi (Eds), *L'évaluation de l'enseignement par les étudiants* (p. 17-34). De Boeck.
- Foucault, M. (2001). « Le jeu de Michel Foucault » dans *Dits et écrits II*, (p. 298-329). Gallimard.
- Fox, D. (1983). Personal theories of teaching. *Studies in Higher Education*, 8, 151-163.
- Freire, P. (1974). *Pédagogie des opprimés - Conscientisation et révolution* (traduit du brésilien). Maspero.
- Fréville, Y. (2001). *Rapport sur la politique de recrutement et la gestion des universitaires et des chercheurs*. Rapport au Sénat.
- Fuller, F.F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207-226.
- Gage, N.L. (1989). The Paradigm Wars and Their Aftermath: A "Historical" Sketch of Research on Teachingsince. *Educational Researcher*, 18, 7, 4-10. <https://www.jstor.org/stable/1177163>.
- Geneva Emotion Research Group. (2002). Questionnaire Genevois d'Appraisal (QGA). Format, développement et utilisation.3.0. [en ligne] : http://www.affectivesciences.org/system/files/webpage/GAQ_Fran%C3%A7ais_0.pdf
- Giddens, A. (1987). *La constitution de la société : Eléments de la théorie de la structuration*, PUF.
- Gingras, Y. (2018). Du mauvais usage de faux indicateurs. *Revue d'histoire Moderne et Contemporaine*, 55 (4bis), 67-79.
- Giroux, H. (1985). Teachers as Transformative Intellectals, *Social Education*, 49, 5, 376-379.
- Glassick, C.E., Huber, M.T. et Maeroff, G. I. (1997). *Scholarship Assessed : Evaluation of the Professoriate*. Jossey-Bass.
- Goasdoué, R. (2013). La folie des grands nombres ou la construction sociale des marchés de la recherche. Dans M. Romainville, R. Goasdoué, R. et M. Vantourout (Dir.), *Évaluation et enseignement supérieur* (p. 189 à 218). De Boeck Supérieur. doi:10.3917/dbu.romai.2013.01.
- Golding, C. et Adam, L. (2016). Evaluate to improve: useful approaches to student evaluation, *Assessment et Evaluation in Higher Education*, 41:1, 1-14, DOI: 10.1080/02602938.2014.976810
- Gori, R. et Del Volgo, M.J. (2009). De la société de la norme à une conception managériale du soin, *Connexions*, 91, 123-147.
- Grandjeand, D. et Scherer, K. (2009). Théorie de l'évaluation cognitive et dynamique des processus émotionnels. In D. Sander et K. Scherer (Ed.), *Traité de psychologie des émotions* (p. 42-76). Dunod.
- Greene, J. C. , et Abma, T. A. (Eds.). (2001). *Responsive evaluation. New directions for evaluation*, No. 92. Jossey-Bass.

- Guba, E. G., et Lincoln, Y. S. (2005). Paradigmatic Controversies, Contradictions, and Emerging Confluences. In N. K. Denzin et Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (p. 191–215). Sage Publications Ltd.
- Guba, E. G., and Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage.
- Guba, E. G., and Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin et Y. S. Lincoln (Eds.), *Handbook of qualitative research* (p. 105–117). Sage Publications, Inc.
- Guba, E.G. and Lincoln, Y.S. (1982). Epistemological and methodological bases of naturalistic inquiry. *ECTJ* 30, 233–252. <https://doi.org/10.1007/BF02765185>.
- Haas, P. M. (1992). *Epistemic Communities and International Policy Coordination, International Organization*, 46/1, 1-35.
- Hadji, C. (1992). *L'évaluation des actions éducatives*. Presses universitaires de France.
- Hall, G., et Loucks, S. (1979). *Implementing innovations in schools: A concerns- based approach*. Austin, TX: Research and Development Center for Teacher Education, University of Texas.
- Harding, S. (1991). *Whose science? Whose knowledge?* Indiana University Press.
- Harding, S. (1998). *Is science multicultural: postcolonialisms, feminisms, and epistemologies*. Indiana University Press.
- Hendry, G. D., Lyon P. M. et Henderson-Smart C. (2007). Teachers' Approaches to Teaching and Responses to Student Evaluation in a Problem-based Medical Program. *Assessment & Evaluation in Higher Education*, 32, 143-157. doi.org/10.1080/02602930600801894
- Heron, J., et Reason, P. (1997). A Participatory Inquiry Paradigm. *Qualitative Inquiry*, 3(3), 274–294. <https://doi.org/10.1177/107780049700300302>
- Heyde, M. et Le Diouris, L. (2009). Trois années d'évaluation des enseignements par les étudiants dans une université de masse. Pourquoi ? Comment ? Et après ? In M. Romainville, et C. Coggi (eds). *L'évaluation de l'enseignement par les étudiants. Approches critiques et pratiques innovantes* (p 57-72). De Boeck.
- Hollinger, R. (1994). *Postmodernism and the social sciences*. Sage.
- Hoyt, D. P. et Lee, E. (2002). IDEA Technical Report No. 12 : Basic data for the revised IDEA system. Manhattan, KS : *The IDEA Center*. [En ligne] <http://www.espacetemps.net/en/articles/jacques-ranciere-les-territoires-de-la-pensee-partagee-en/>.
- Huberman, M. (1989). Les phases de la carrière enseignante : un essai de description et de prévision. *Revue française de pédagogie*, 80, 5-16.
- Husserl, E. (1947). *Méditations cartésiennes*, trad. fr. par G. Pfeiffer et E. Levinas, Paris : Vrin.
- Husserl, E. (1996). *Leçons pour une phénoménologie de la conscience intime du temps*, trad. H. Dussort. PUF.

- Hutchings, P. et Shulman, L. S. (1999). The Scholarship of Teaching : New Elaborations, New Developments. *Change*, 31 (5), 10-15.
- Hutchings, P., Huber, M. et Ciccone, A. (2011). *The Scholarship of Teaching and Learning Reconsidered*. Jossey-Bass.
- Joas, H. (1999). *La créativité de l'agir*. Traduction Pierre Rusch. CERF.
- Johnson, T. (1999). Course Experience Questionnaire 1998. Parkville, victoria: Graduate carrers concil of australia.
- Jorro, A. (2007). *Évaluation et développement professionnel*. L'Harmattan.
- Josephs, R. A., Larrick, R. P., Steele, C. M., et Nisbett, R. E. (1992). Protecting the self from the negative consequences of risky decisions. *Journal of Personality and Social Psychology*, 62(1), 26–37. <https://doi.org/10.1037/0022-3514.62.1.26>.
- Karsenti, T. (2014). Modèle ASPID. Récupéré du site de l'auteur : <http://karsenti.ca>
- Katz, L. (1972). Developmental stages of preschoolteachers. *Elementary School Journal*, 73, 50- 55.
- Kember, D. et McNaught, C. (2007). *Enhancing University Teaching*. Routledge.
- Kim, L.E., Jörg, V. et Klassen, R.M. (2019). A Meta-Analysis of the Effects of Teacher Personality on Teacher Effectiveness and Burnout. *Educ Psychol Rev* 31, 163–195. <https://doi.org/10.1007/s10648-018-9458-2>
- Kogan, L. R., Schoenfeld-Tacher, R. et Hellyer, P. W. (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education*, 15(6), 623-636.
- Kolitch, E. and Dean, A. (1999). Student Ratings of Instruction in the USA: Hidden Assumptions and Missing Conceptions About 'Good' Teaching”, *Studies in Higher Education*, 24(1), 27-43.
- Kreber, C. et Cranton, P. (2000). Exploring the Scholarship of Teaching, *Journal of Higher Education*, 71, (4), 476-495.
- Kuhn, T. (1990). La tension essentielle (traduction française de Second Thoughts on Paradigms, in *The Essential Tension, Selected Studies in Scientific Tradition and Change*, 1977, The University of Chicago Press). Gallimard.
- Kuhn, T.S. (1983). *La structure des révolutions scientifiques* (traduit par L. Meyer, The Structure of Scientific Revolutions paru en 1962 et 1970, première traduction française 1972). Flammarion.
- Laoureux, S., et Frère, B. (2013). Promesses et impasses sociologiques de la phénoménologie. In *La phénoménologie à l'épreuve des sciences humaines* (p. 135-157). Peter Lang.
- Lapeyronnie, D. (2004). L'académisme radical ou le monologue sociologique: Avec qui parlent les sociologues ? *Revue française de sociologie*, 45(4), 621-651. doi:10.3917/rfs.454.0621.

- Lather, P. (1986). Issues of validity in openly ideological research: Between a rock and a soft place. *Interchange* 17, 63–84, <https://doi.org/10.1007/BF01807017>.
- Lather, P. (1993). Fertile Obsession: Validity after Poststructuralism. *The Sociological Quarterly*, 34(4), 673-693. Retrieved June 1, 2020, from www.jstor.org/stable/4121374
- Lather, P. (2006). Paradigm proliferation as a good thing to think with: teaching research in education as a wild profusion, *International Journal of Qualitative Studies in Education*, 19:1, 35-57, DOI: 10.1080/09518390500450144.
- Laugier, S. (2003). Signification et incommensurabilité : Kuhn, Carnap, Quine. *Archives de Philosophie*, tome 66(3), 481-503. doi:10.3917/aphi.664.0481.
- Laveault, D. (2007). De la «régulation» au «réglage»: élaboration d'un modèle d'autoévaluation des apprentissages. Dans L. Allal et L. Mottier Lopez (Dir.), *Régulation des apprentissages en situation scolaire et en formation*. De Boeck.
- Laveault, D. (2012). Autorégulation et évaluation-soutien d'apprentissage. Dans L. Mottier Lopez et G. Figari (Dir.), *Modélisations de l'évaluation en éducation. Questionnements épistémologiques* (p. 115-130). De Boeck.
- Lazarus, R. S. (1966). *Psychological stress and the coping process*. McGraw Hill.
- Lebrun, M. (2011). Dispositif hybride et flipped classrooms. Blog de M@rcel. Retrieved from: <http://lebrunremy.be/WordPress/?p=583>.
- Leclercq, B. (2012). La valeur du savoir, dans JE Charlier, S. Croché et B. Leclercq (dir.) *Contrôler la qualité dans m'enseignement supérieur*, (p.135-168), Academia.
- Lepenes, W. (1990). *Les trois cultures. Entre science et littérature, l'avènement de la sociologie* (traduit par H. Plard). Editions de la Maison des sciences de l'homme.
- Lévy, J., Rennes, J. et Zerbib, D. (2007). Jacques Rancière : « Les territoires de la pensée partagée ». *EspacesTemps.net, Laboratory*, 08.01.2007. [En ligne] <https://www.espacestems.net/articles/jacques-ranciere-les-territoires-de-la-pensee-partagee/>.
- Lewin, K. (1959). *Psychologie dynamique*. Presses universitaires de France.
- Lincoln, Y. S. (1995). Emerging Criteria for Quality in Qualitative and Interpretive Research. *Qualitative Inquiry*, 1 (3), 275-28.
- Lincoln, Y. S. (1998). From Understanding to Action: New Imperatives, New Criteria, New Methods for Interpretive Researchers, *Theory et Research in Social Education*, 26,1, 12-29, DOI: 10.1080/00933104.1998.10505832
- Lincoln, Y. S. et Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Lincoln, Y. S., et Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin, et Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 1065-1122), Thousand Oaks, CA: Sage Publications.

- Lordon, F. (1997). Le désir de "faire science". *Actes de la recherche en sciences sociales*, 119, 27-35. doi : <https://doi.org/10.3406/arss.1997.3227>
- MacGeorge, E., Homan, S., Dunning, J., Elmore, D., Bodie, G., Evans, E., Khichadia, S., Lichti, S., Feng, B., et Geddes, B. (2008). Student evaluation of audience response technology in large lecture classes. *Educational Technology Research et Development*, 56(2), 125-145.
- Machell, D. F. (1989). A discourse on professional melancholia. *Community Review*, 9(1-2), 41-50.
- Madaus, G.F., Scriven, M., et Stufflebeam, D.L. (1983). *Evaluation models: Viewpoints on educational and human services evaluation*. Kluwer Nijhoff.
- Madaus, G.F., Stufflebeam, D.L., et Kellaghan, T. (2000). *Evaluation models : Viewpoints on educational and human services evaluation*. (2nd ed.). Kluwer Academic Publishers.
- Marcel, J.-F. (2009). De la prise en compte des pratiques enseignantes de travail partagé. *Nouveaux cahiers de la recherche en éducation*, 12 (1), 47-64. <https://doi.org/10.7202/1017487ar>.
- Marcel, J.-F. et Murillo, A. (2014). Analyse du fonctionnement de collectifs d'enseignants : proposition méthodologique. *Questions Vives* [En ligne], n° 21. DOI : <https://doi-org.ezproxy.uca.fr/10.4000/questionsvives.1507>.
- Marcel, J.-F., et Broussal, D. (2017). *Emancipation et recherche en éducation. Conditions de la rencontre entre science et militance*. Editions du Croquant.
- Marsh, H. W. et Ware, J. E. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student ratings scales. New interpretations of the Dr. Fox effect. *Journal of Educational Psychology*, 74, 126-134.
- Marsh, H. W. (1987). Students' evaluations of university teaching : Research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11 (3), 255-379.
- Marsh, H. W., Fleiner, H. et Thomas, C. S. (1975). Validity and Usefulness of Student Evaluations of Instructional Quality. *Journal of Educational Psychology*, 67, 833-839.
- Marsh, H.-W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology*, 52, 77-95.
- Marsh, P. A. (2007). What is Known about Student Learning Outcomes and How does it relate to the Scholarship of Teaching and Learning ? *International Journal for the Scholarship of Teaching and Learning*, 1 (2). [En ligne] <https://doi.org/10.20429/ijsotl.2007.010222>.
- Martinot, D. (2016). Le soi en psychologie sociale. Dans : Catherine Halpern éd., *Identité(s): L'individu, le groupe, la société* (pp. 37-45). Auxerre, France: Editions Sciences Humaines. doi:10.3917/sh.halpe.2016.01.0037.
- Masterman, M. (1970). The Nature of a Paradigm. In I. Lakatos and A. Musgrave (Eds.), *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy*

of Science, London, 1965 (pp. 59-90). Cambridge University Press.
doi:10.1017/CBO9781139171434.008

- Mc Murtry, J. (1991). Education and the market model. *Journal of Philosophy of Education*, 25, 209-217.
- McKeachie, W. J. (1979). Student rating of faculty : A reprise. *Academe*, 65, 384-397.
- McKenzie, J. (1995). Bridges and barriers : factors influencing changes in conceptions of teaching. A pilot study. *Research and Development in Higher Education*, 18, 533-538.
- Mead, G. H. (1963). *L'esprit, Le soi et la société* (1934), traduction de J. Cazeneuve, E. Kaelin et J. Thibault. PUF.
- Mertens, D. M. (2015). Philosophical assumptions and program evaluation. *Spazio Filosofic*, 13, 75-85.
- Mertens, D. M., and Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. Guilford.
- Mertens, D. M. (2009). *Transformative Research and Evaluation*. Guilford.
- Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, 18(2), 5–11. <https://doi.org/10.3102/0013189X018002005>.
- Meyer, M. et Molyneux-Hodgson, S. (2011). « Communautés épistémiques » : une notion utile pour théoriser les collectifs en sciences ? *Terrains et travaux*, 18(1), 141-154. <https://www.cairn.info/revue-terrains-et-travaux-2011-1-page-141.htm>.
- Mialaret, G. (2011). *Les sciences de l'éducation* (1^{re} éd. 1976). PUF, Collection « Que sais-je ? ».
- Miesen, M. van der. (2015). The Role of Self-Esteem in Social Feedback : An fMRI Study. *Student Undergraduate Research E-Journal!*, 1(1).
- Mishra, P. et Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teacher knowledge. *Teachers College Record*, 108(6), 1017-1054.
- Moersch, C. (2001). Next steps: Using LoTi as a research tool. *Learning et Leading With Technology*, 29 (3), 22-27.
- Moore, S., et Kuol, N. (2005). Students evaluating teachers: exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education* 10(1).
- Morais, M.A. (2001). *Les 5 niveaux d'appropriation des technologies de l'information et de la communication chez les enseignantes et les enseignants*. Shédiac, N.-B. : District scolaire n° 1.
- Moredich, C., et Moore, E., (2007). Engaging students through the use of classroom response systems. *Nurse Educator*, 32(3), 113-116.

- Morrisette, J., Mottier Lopez, L. et Tessaro, W. (2012). La production de savoirs négociés dans deux recherches collaboratives sur l'évaluation formative. In L. Mottier Lopez et G. Figari (Ed.), *Modélisations de l'évaluation en éducation: questionnements épistémologiques* (p. 27-43). De Boeck.
- Moscovici, S. (1984). *Psychologie sociale*. Presses Universitaires de France.
- Moscovici, S. et Paicheler, G. (1973). Travail, individu et groupe, in S. Moscovici (dir.), *Introduction à la psychologie sociale*, t. 2, (p. 9-44). Larousse.
- Moser, G. (2006). Psychologies sociales. Psychologie sociale, application de la psychologie sociale et psychologie sociale appliquée. *Les Cahiers Internationaux de Psychologie Sociale*, 70(2), 89-95. doi:10.3917/cips.070.0089.
- Moses, I. (1988). *Academic Staff Evaluation and Development: a university case study*. University of Queensland Press.
- Moskal, A. C. M., Stein, S. J., et Golding, C. (2016). Can You Increase Teacher Engagement with Evaluation Simply by Improving the Evaluation System? *Assessment and Evaluation in Higher Education*, 41(2), 286-300.
- Mottier Lopez, L. (2008). *Apprentissage situé: la microculture de classe en mathématiques*. Berne: Peter Lang.
- Mottier Lopez, L. et Figari, G. (Eds.) (2012). *Modélisation de l'évaluation en éducation. Questionnements épistémologiques*. De Boeck.
- Mottier Lopez, L., Borloz, S., Grimm, K., Gros, B., Herbert, C., Methenitis, J., Payot, C., Pot, M., Schneuwly, J.-L. et Zbinden, S. (2010). Les interactions de la régulation entre l'enseignant et ses élèves. Expérience d'une recherche collaborative. Dans L. Lafortune, S. Fréchette, N. Sorin, P.-A. Doudin et O. Albanese (Ed.), *Approches affectives, métacognitives et cognitives de la compréhension* (p. 33-50). Presses de l'Université du Québec.
- Murray, H. G. (1984). The Impact of Formative and Summative Evaluation of Teaching in North American Universities. *Assessment and Evaluation in Higher Education*, 9 (2), 117-132.
- Musselin, C. (2008). *Les universitaires*. La Découverte.
- Nancy, J.-L. (1999). *La communauté désœuvrée*. Christian Bourgeois.
- Nasser, F., et Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education*, 27(2), 187-198.
- Nault, T. (1999a). Écllosion d'un moi professionnel personnalisé. Dans J.-C. Héту, M. Lavoie et S. Baillauquès (dir.), *Jeunes enseignants et insertion professionnelle* (p. 139-159). De Boeck.
- Nault, T. (1999b). Les forces d'incubation pour un moi professionnel personnalisé en enseignement. Dans J.-C. Héту, M. Lavoie et S. Baillauquès (dir.), *Jeunes enseignants et insertion professionnelle* (p. 139-159). De Boeck.

- Norris, J., et Conn, C. (2005). Investigating Strategies for Increasing Student Response Rates to Online-Delivered Course Evaluations. *Quarterly Review of Distance Education*, 6, 1, 13-29.
- Nulty, D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment et Evaluation in Higher Education*, 33, 3, 301-314.
- Overall, J. U. et Marsh, H. W. (1979). Midterm Feedback from Students: Its Relationship to Instructional Improvement and Students' Cognitive and Affective Outcomes. *Journal of Educational Psychology*, 71, 856-865.
- Paivandi, S. et Younès, N. (2019). *À l'épreuve d'enseigner à l'Université. Enquête en France*. Berne : Peter Lang.
- Paivandi, S. et Younès, N. (2017). L'EEE et la régulation des pratiques pédagogiques à l'université en France. *Éducation et formation*, 307.
- Pajares, M. F. (1992). Teachers' beliefs and educational research : cleaning up a messy construct. *Review of Educational Research*, 62 (3), 307-332.
- Patton, M.Q. (1980). *Qualitative evaluation methods*. Sage.
- Peeters, H. et Charlier, P. (1999). Contributions à une théorie du dispositif. *Hermès, La Revue*, 25(3), 15-23. <https://www.cairn.info/revue-hermes-la-revue-1999-3-page-15.htm>.
- Perret, C. (2017). L'évaluation des enseignements par les étudiants peut-elle participer au développement professionnel pédagogique des enseignants de l'université française ? *Éducation et formation*, 307 (1). [En ligne].
- Polanyi, M. (1958). *Personal Knowledge, Towards a Post-Critical Philosophy*. Routledge et Kegan Paul.
- Postiaux, N. (2010). *Référentiels de compétences et pilotage de formation à l'université : rôle, enjeux et limites*. Thèse de doctorat en sciences psychologiques et de l'éducation, Université libre de Bruxelles.
- Puenteadura, R. (2009) TPCK and SAMR Models for Enhancing technology Integration (Transcript), As we may Teach: educational Technology, From Theory Into Practice. Retrieved from: <https://itunes.apple.com/itunes-u/as-we-may-teach-educational/id380294705?mt=10#ls=1>.
- Puenteadura, R. (2014). SAMR in the classroom [billet de blogue]. Récupéré du blogue de l'auteur : <http://hippasus.com/rrpweblog>
- Puenteadura, R. (2014). SAMR In The Classroom. Retrieved from: <http://www.hippasus.com/rrpweblog/archives/2014/08/27/SAMRInTheClassroom.pdf>.
- Py, J., Somat, A. et Baillé, J. (1999). *Psychologie sociale et formation professionnelle*. Presses Universitaires de Rennes.
- Quinton, A. (2009). L'évaluation interne des enseignements et des formations : un débat entre pairs avec la participation des étudiants. In M. Romainville, et C. Coggi (eds). *L'évaluation de*

l'enseignement par les étudiants. Approches critiques et pratiques innovantes (p 73-93). De Boeck.

- Raby, C. (2004). *Analyse du cheminement qui a mené des enseignants du primaire à développer une utilisation exemplaire des technologies de l'information et de la communication (TIC) en classe*. Thèse de doctorat. Université du Québec à Montréal, Montréal. Retrieved from: <https://tel.archives-ouvertes.fr/edutice-00000750>.
- Ramsden, P. (1992). *Learning to teach in higher education*. Kogan Page.
- Rancière, J. (1995). *La méésentente. Politique et philosophie*. Galilée.
- Rancière, J. (2000). *Le partage du sensible. Esthétique et politique*. La Fabrique.
- Rege Colet, N. (2009). L'évaluation de l'enseignement au cœur des processus d'assurance qualité : l'arbre qui cache la forêt. Dans M. Romainville et C. Coggi (Dir.), *L'évaluation de l'enseignement par les étudiants: Approches critiques et pratiques innovantes* (p. 235-253). De Boeck Supérieur. doi:10.3917/dbu.romai.2009.01.0235.
- Reinharz, S. (1997). Who am I ? The need for a variety of selves in the field. In R. Hertz (Ed.), *Reflexivity and voice*, p. 3–20. Sage.
- Rey, A. (dir.) (1994). *Dictionnaire historique de la langue française*. Dictionnaire Le Robert.
- Richardot, S. (2006). Regards sur la psychologie sociale. *Les Cahiers Internationaux de Psychologie Sociale*, 71(3), 41-53. doi:10.3917/cips.071.0041.
- Richardson, L. (1994). Writing: A method of inquiry. In N. K. Denzin et Y. S. Lincoln (Eds.), *Handbook of qualitative research* (p. 516–529). Sage Publications, Inc.
- Richardson, L. (1997). *Fields of Play : Constructing an Academic Life*. Rutgers University Press.
- Richardson, L. (1997). *Fields of play: Constructing an academic life*. Rutgers University Press.
- Ritzer, G. (1996). McUniversity in the postmodern consumer society. *Quality in Higher Education*, 2 (3), 185-199.
- Rodger, S., Murray, H.G., Cummings, A.L. (2007). Effects of teacher clarity and student anxiety on student outcomes. *Teaching in Higher Education* 12(1):91-104. DOI: 10.1080/13562510601102255.
- Romainville, M. (2009). Une expérience d'élaboration collective de critères de qualité. Dans M. Romainville et C. Coggi (dir.), *L'évaluation de l'enseignement par les étudiants. Approches critiques et pratiques innovantes* (p. 145-163). De Boeck.
- Romainville, M. et Coggi C. (2009). *L'évaluation de l'enseignement par les étudiants*. De Boeck.
- Rorty, R. (2017). *La Philosophie et le Miroir de la nature* (traduit par T. Marchaisse, 1ere éd.1979). Seuil.
- Rosa H. (2010). *Accélération. Critique sociale du temps*. La Découverte.

- Rowley, J. (1996). Measuring quality in higher education. *Quality in Higher Education* 2 (3), 237-255.
- Salcin, A., Wéry, F., Emplit, P. et Robert, F. (2012). Les perceptions des enseignants-chercheurs de l'Université libre de Bruxelles à propos de l'évaluation des enseignements. *Mesure et évaluation en éducation*, 35 (3), 61-84.
- Sandholtz, J. H., Ringstaff, C., et Owyer, D. C (1997). *La classe branchée: Enseigner à l'ère des technologies*. Chenelière/McGraw-Hill.
- Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr et T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, Methods, Research* (p. 92-120). Oxford University Press.
- Schmelkin, L. P., Spencer, K. J., et Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education*, 38(5), 575-592.
- Schmelkin, L.P., Spencer, K.J. et Gellman, E.S. Faculty Perspectives on Course and Teacher Evaluations. *Research in Higher Education* 38, 575-592 (1997). <https://doi.org/10.1023/A:1024996413417>
- Schmidt, B. and Loßnitzer, T. (2010) Lehrveranstaltungs evaluation: State of the Art, ein Definitionsvorschlag und Entwicklungslinien, *Zeitschrift für Evaluation*, 9.1, 49-72.
- Schnoz-Schmied, T. P. (2017). Vers un soutien ciblé au développement de la formation. *Education et Formation*, n°e-307-01, 145-166.
- Schön, D. (1994). *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel* (J. Heynemand et D. Gagnon, trad.). Les Éditions Logiques. (Ouvrage original publié en 1983 sous le titre *The Reflective Practitioner*. Basic Books.
- Schütz, A. (1962). *Collected Papers : Tome I : The Problem of Social Reality*. Éditions Martinus Nijhoff.
- Schütz, A. (1987). *Les chercheurs et le quotidien*. Editions Méridiens-Klincksieck.
- Schwandt, T. A. (1996). Farewell to criteriology. *Qualitative Inquiry*, 2(1), 58-72. <https://doi.org/10.1177/107780049600200109>.
- Scriven, M. (1967). The methodology of evaluation, in Tyler, R. ; Gagne, R. et Scriven, M. (dir.) (1967). *Perspectives of curriculum evaluation*, AERA Monograph Series on Curriculum Evaluation, No. 1 (p. 39 83). Rand McBally.
- Scriven, M. (1980). *The logic of evaluation*. Edgepress.
- Scriven, M. (1983). Evaluation ideologies. In Madaus, G.F., Scriven, M., et Stufflebeam, D.L. (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation*. Kluwer Nijhoff.
- Scriven, M. (1986). New frontiers of evaluation, *Evaluation Practice*, 7, 7-44.

- Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*, 40 (1), A40.
- Shadish, W.R., Cook, T.D. et Leviton, L.C. (1991). *Foundations of Program Evaluation*. SAGE.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. Récupéré du site CiteSeerX : <http://citeseerx.ist.psu.edu>
- Shulman, L. S. (1993). Teaching as community property : Putting an end to pedagogical solitude. *Change*, 25 (6), 6-7.
- Smith, R. (2001). Formative Evaluation and the Scholarship of Teaching and Learning. In C. Knapper and P. Cranton (Eds.), *Fresh Approaches to the Evaluation of Teaching, New Directions for Teaching and Learning*, n° 88, p. 51- 62.
- Spiller, D., et Harris, T. (2013). Learning from evaluations: Probing the reality. *Issues in Educational Research*, 23(2 SPL), 258–268.
- Sproule, R., et Vâlsan, C. (2009). The student evaluation of teaching: its failure as a research program, and as an administrative guide. *The AMFITEATRU ECONOMIC journal, Academy of Economic Studies*, 11(25), 125-150.
- Stake, R. E. (1975). To evaluate an arts program. In R. E. Stake , *Evaluating the arts in education: A responsive approach* (p. 13-31). Merrill.
- Stake, R.E. (1976). A Theoretical Statement of responsive Evaluation. *Studies in Educational Evaluation*, 2, 19-22.
- Stengers, I. (2006). *La Vierge et le Neutrino. Les scientifiques dans la tourmente*, Les Empêcheurs de penser en rond.
- Stufflebeams, L. et Shinkfield, A. J. (1985). *Systematic Evaluation*. Kluwen et Nijhoff.
- Tashakkori, A. et Teddlie, C. (2003). *Handbook of mixed-methods in social et behavioral research*. Sage.
- Teddlie, C. et Tashakkori, A. (2009). *Foundations of Mixed Methods Research. Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage.
- Theall, M. (2002). Student rating : Myths vs research evidence. *Brigham Young University's FOCUS ON FACULTY Newsletter*, 10(3), 2.
- Theall, M. et Franklin, J. (2001). Using Technology to Facilitate Evaluation. *New Directions for Teaching and Learning*, 88, 41-50.
- Touraine, A. (1994). *Qu'est-ce que la démocratie ?* Fayard.
- Trees, A. and Jackson, M. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, 32(1), 21-40.

- Trigwell, K. et Prosser, M. (1996). Changing approaches to teaching : a relational perspective, *Studies in Higher Education*, 21 (3), 275-284.
- Trowler, P., Saunders, M. et Bamber, V. (2012). *Tribes and Territories in the 21st Century : Rethinking the significance of disciplines in higher education*. Routledge-Open University.
- Van Damme, D. (2003). Standards and Indicators in Institutionnal and Program Accreditation in Higher Education. Rapport UNESCO-CEPES
- Vergnaud, G. (1996). Au fond de l'action, la conceptualisation. Dans J.-M. Barbier (dir.), *Savoirs théoriques et savoirs d'action* (p. 275-292). Presses universitaires de France.
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation, méthodes dispositifs, outils*. De Boeck.
- Vinokur, A. (2006). La qualité de la mesure de la qualité dans l'enseignement supérieur : essai d'analyse économique. *Éducation et sociétés*, 18(2), 109-124. doi:10.3917/es.018.0109.
- Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education* 23, 2, 191-211.
- Weimer, M., Menges, R.J. and associates (1996). *Teaching on Solid Ground : Using Scholarship to Improve Practice*. Jossey-Bass.
- Weiss, C., and Bucuvalas, M. (1980). *Social Science Research and Decision-Making*. Columbia University Press.
- Weiss, C.H. (1987). Evaluating social programs : What we have learned ? *Society*, 25, 1, 40-45.
- Wenger, E. (2000). Communities of Practice and Social Learning Systems, *Organization*, 7/2, 225-246.
- Wholey, J.S. (1983). *Evaluation and public effective management*. Little Brown.
- Willett, G. (2012). Paradigme, théorie, modèle, schéma : qu'est-ce donc ? *Communication et organisation*, 10 [Online], DOI : 10.4000/communicationorganisation.1873
- Yao, Y. et Grady, M. (2005). How Do Faculty Make Formative Use of Student Evaluation Feedback? A Multiple Case Study. *Journal of Personnel Evaluation in Education*, 18(2), 107-126.
- Yin, R. K. (1984). *Case study research : Design and methods*. Sage.
- Younes, N. et Romainville, M. (2012). Les transformations actuelles de l'évaluation de l'enseignement par les étudiants. *Mesure et évaluation en éducation*, 35, 3. 177-201.
- Younès, N ; Rege Colet, N.; Detroz, P. et Sylvestre, E. (2013). L'évaluation de l'enseignement par les étudiants, une dynamique paradoxale. In M. Romainville, R. Goasdoué et M. Vantourout, (Dir.). *Evaluation et enseignement supérieur*. De Boeck.
- Younes, N. (2005). Vers une nouvelle culture de l'enseignement universitaire ? L'évaluation de l'enseignement par les étudiants avec les TIC, *Les Dossiers des Sciences de l'Éducation*, 13, 67-94.

- Younès, N. (2006). *L'effet évaluation de l'enseignement supérieur par les étudiants*. Thèse de doctorat non publiée, Université Pierre Mendès France, Grenoble.
- Younès, N. (2007). À quelles conditions l'évaluation formative de l'enseignement par les étudiants est-elle possible en France ? *Revue française de pédagogie*, 161, 25-40.
- Younès, N. (2009). Multidimensionnalité de l'évaluation de l'enseignement universitaire dans la littérature anglophone. Dans V. Bedin (dir.), *L'évaluation à l'université. Évaluer ou conseiller ?* (p. 191-205). Presses universitaires de Rennes.
- Younès, N. (2013). Perspective d'évaluation « écologique » des dispositifs éducatifs et de formation. In V. Bedin et L. Talbot (Dir.), *Les points aveugles dans l'évaluation des dispositifs d'éducation ou de formation* (p. 71-87). Peter Lang.
- Younès, N. (2015). L'évaluation de l'enseignement par les étudiants : le tournant du partage. *Évaluer. Journal international de recherche en éducation et formation*, 1, 79-90. [En ligne] http://www.e-jref.org/docannexe/file/92/e_jref_1_1_younes_79_90.pdf.
- Younès, N. (2018, juillet). *Auto-évaluation and réflexivité dans l'enseignement supérieur*. Université d'été de l'ADMEE-Europe, Balamand, Liban.
- Younès, N. et Gay, J.C. (2014). L'évaluation des enseignements par les étudiants dans les démarches qualité : vers une alternative à des dérives quantitatives normatives et consuméristes. In C. Fallon et B. Leclercq (Dir). *Leurres de la qualité dans l'enseignement supérieur? Variations internationales sur un thème ambigu* (p. 203-216). Academia.
- Younès, N. et Paivandi, S. (2017). Expérimentation d'un dispositif coopératif d'EEE entre enseignants et chercheur-médiateur dans le cadre d'un master de formation d'enseignants. *Éducation et formation*, 307 (1), 167-182. [En ligne] <http://revueeducationformation.be/>.
- Younès, N., Gremion, C. et Sylvestre, E. (2020). *L'évaluation sources de synergies ? Entre normalisation, contrôle et développement formatif*. Presses de l'ADMEE.
- Younès, N., Sasse, M. et Darj, J.M. (2016). L'enjeu de la subjectivation dans les pratiques d'évaluation à l'école élémentaire. In L. Mottier Lopez et W. Tessaro (Eds), *Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages* (p.149-167). Peter Lang.
- Younès, N., Soulier, S. et Detroz, P. (2017). Utilisation des boîtiers de vote et développement professionnel. *Revue internationale des technologies en pédagogie universitaire*, 14 (3), 17. [En ligne] <https://doi.org/10.18162/ritpu-2017-v14n3-02>.
- Zabaleta, F. (2007). the use and misuse of student evaluation of teaching. *Teaching in Higher Education*, 12(1), 55-76. doi:10.1080/13562510601102131